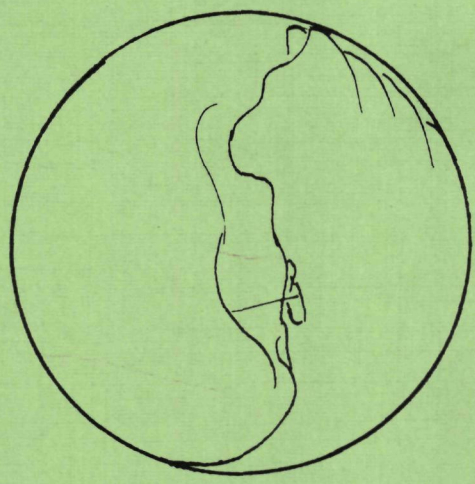
Characterization of the human and feline c-*sis* proto-oncogenes



Ans van den Ouweland

Characterization of the human and and feline c-*sis* proto-oncogenes

sponsored by Amersham

Aan mijn moeder en Lambert

De totstandkoming van een proefschrift vergt de inzet van veel personen. Daarom wil ik ook een ieder die iets bijgedragen heeft daarvoor hartelijk bedanken.

Allereerst mijn moeder en Lambert, die een grote steun en toeverlaat zijn geweest gedurende mijn gehele studietijd.

Peter Bloemers en Wim van de Ven voor hun prettige begeleiding en de grote vrijheid die ik tijdens mijn onderzoek genoten heb.

Jan van Groningen die de laatste loodjes op een voortreffelijke wijze mee heeft helpen dragen.

De andere leden van de Molonc groep: Jack Schalken, Carla Onnekink, Anton Roebroek, Han van Neck, Cor Lamers, Maddy de Jong, Adrie van Bokhoven en Sjef Verbeek voor hun prettige samenwerking gedurende dit onderzoek.

Marco Breuer, Inge van Renesse en Peter Hendriksen die in het kader van hun doctoraalstudie betrokken waren bij dit onderzoek.

Verder alle andere leden van de afdeling Biochemie voor de uitstekende werksfeer.

De Nederlandse belastingbetalers voor hun financiële steun.

Characterization of the human and feline *c-sis* proto-oncogenes

Proefschrift

ter verkrijging van de graad van doctor
in de wiskunde en natuurwetenschappen
aan de Katholieke Universiteit te Nijmegen
op gezag van de Rector Magnificus
Prof. Dr. B.M.F. van Iersel
volgens besluit van het College van Decanen
in het openbaar te verdedigen
op vrijdag 12 juni 1987
des namiddags te 1.30 precies.

door

Anna Maria Wilhelmina van den Ouweland

Geboren te Tilburg

Promotor : Prof. Dr. H.P.J. Bloemers

Co-referent: Dr. W.J.M. van de Ven

CONTENTS

CHAPTER 1

INTRODUCTION

Retroviruses and proto-oncogenes.

Retroviruses have been isolated from a wide variety of animal species. In their natural hosts, they produce a number of pathological syndromes including malignant tumors, fatal proliferative diseases (not necessarily neoplastic), anemias and slow degenerative diseases that affect hemapoietic, neural and other tissues. Tumors induced by retroviruses can be classified as sarcomas, carcinomas, leukemias and lymphomas. As far as tumor induction is concerned, retroviruses may be divided into two classes: acutely and nonacutely transforming retroviruses. The acutely transforming viruses cause the appearance of a tumor within a few weeks after infection. The nonacutely transforming viruses induce tumors slowly; "slow" referring to the long latency period between infection and the appearance of the disease.

The nonacutely transforming retroviruses are replication competent and possess an RNA genome that contains three different genes: gag, encoding proteins of the viral nucleoid, pol encoding reverse transcriptase and env, which encodes proteins of the viral envelope. The latter are also present in the host cell membrane. All virus-encoded products are required for viral replication. The nonacutely transforming retroviruses may induce various types of malignancies although they have not acquired a so-called oncogene (i.e. a gene directly responsible for tumorigenesis and often for transformation of cultured cells). For a long time the pathogenic properties of these viruses were enigmatic. After a period of confusing theories

and conflicting interpretations the following conclusions are now generally accepted.

A virus particle is able to infect cells which express a particular type of receptor protein on their cell surface that is recognized by the main viral *env* protein, such as gp70. Therefore, gp70 determines the organ tropism of a virus strain. In infected animals, the organ tropism may undergo changes when new gp70 molecules arise as a result of recombinations between the pathogenic virus and so-called endogenous usually nonpathogenic viruses.

In each newly infected cell, the viral genome is integrated somewhere in the cellular genome. Viral regulatory sequences, especially viral promoters and enhancers, may influence the expression of genes in the neighborhood of the integration site (promoter or enhancer insertion). This is called cis-activation, because it is limited to the chromosome that contains the integrated viral element. When, in this way, a proto-oncogene (see below) is activated, the cell may become a malignant cell. The long and variable periods of latency reflect the fact that such a malignant activation is a rare event in any one cell but likely to happen sometimes in an organism. Alternatively, integrated viruses may activate cellular genes through so-called trans-activation, *i.e.* via a soluble factor whose action is not limited to the site of integration. The mechanism is not yet well understood. The pathogenicity of HIV for example is thought to involve such trans-activation through the action of the *tat* gene (trans-activating transcriptional gene) of the virus (Wong-Staal and Gallo, 1985; Greene *et al*, 1986).

The acutely transforming retroviruses lead to transformation within a relatively short period after infection. In contrast to nonacutely transforming retroviruses, they transform cells both *in vitro* and *in vivo*. The initiation and maintenance of neoplastic transformation induced by acutely transforming viruses are caused by the products of viral oncogenes present in these viruses. In almost all cases, the oncogenes replace viral genes necessary for replication, with the consequence that the acutely transforming retroviruses are replication defective and need helper virus for their propagation. The viral oncogene (v-*onc*) of the acutely transforming retroviruses is the result of a recombination between the genomes of nonacutely transforming retroviruses and proto-oncogenes (c-*onc*) (Fishinger, 1982; Bishop, 1983). As a consequence of this recombination, the transforming potential of proto-oncogenes is mobilized. Apparently, such recombination events have taken place in one of the many tumor cells after cis-activation of a cellular proto-oncogene by a nonacutely transforming retrovirus.

Because of the effect of viral oncogenes on regulation of cell growth and differentiation and because of the conservation of oncogene sequences during evolution, it is reasonable to assume that proto-oncogenes have important functions in normal cellular processes. The normal cellular functions of the proto-oncogenes may sometimes be disturbed. This can be caused by somatic changes within these genes such as point mutations, translocations, gene amplifications or insertions. Interestingly, such genetic changes are frequently found in naturally occuring malignancies suggesting a relationship with these diseases. For an extensive review on retroviruses, viral on-

cogenes and proto-oncogenes see: R. Weiss et al (1984).

With respect to their biological activity, the proto-oncogene products can be divided into several classes such as: cytoplasmic tyrosine phosphokinases (e.g. c-abl, c-fes), potential kinases (e.g. c-mos, c-raf), GTP-binding proteins (c-ras), DNA-binding proteins (e.g. c-myc), growth factors (c-sis) or receptors (c-erbA, c-erbB and c-fms). For the products of some of the proto-oncogenes, the biological activity is not established yet (e.g. c-rel) (Hunter, 1984). The observation that proto-oncogene products constitute or resemble growth factors and growth factor receptors is of particular interest. The relation between c-sis and a growth factor was established when the deduced amino-acid sequence of the v-sis oncogene was shown to exhibit a high degree of homology with the incomplete sequence of platelet-derived growth factor chain 2 (PDGF-2) (Deuel et al, 1983; Doolittle et al, 1983; Waterfield et al, 1983). Other studies revealed that the gene products of the (proto-) oncogenes, c-fms, c-erbA and v-erbB constitute or resemble the receptor for colony stimulating factor-1 (CSF-1) (Sherr et al, 1985), thyroid hormone (Sap et al, 1986; Weinberger et al, 1986) and epidermal growth factor (EGF) (Downward et al, 1984), respectively. These examples suggest that the products of proto-oncogenes are integrated components of signal transduction systems and as such play an important role in processes that control cell growth and differentiation. It is, therefore, not difficult to imagine that their malfunctioning could lead to derailment of normal cell proliferation and tumorigenesis. In the studies described in this thesis, attention is focussed upon the c-sis proto-oncogene.

Growth factors and tumorigenesis.

Part of our knowledge on cell growth *in vivo* and the signals controlling these processes has been derived from studies on established cell lines. Often these cultured cells require serum in their media for their propagation. Serum can be replaced by specific polypeptides, called growth factors (Barnes and Sato, 1980). Growth factors regulate cell proliferation through binding to specific cell membrane receptors (Heldin and Westermark, 1984). They are present in a variety of tissues and are released by many cells in culture (Massagué, 1985; Nilsson *et al*, 1985; Rall *et al*, 1985). The fine tuning of proliferation rates necessary for coordinated growth of various cell types to form and maintain tissues might be due to the great diversification of growth factors, the receptor dependent cell type specific action of these factors and the requirement of multiple growth factors for stimulation of a specific cell (Barnes and Sato, 1980; Walthall and Ham, 1981; Tsao *et al*, 1982; Heldin and Westermark, 1984; Weinberg, 1985; Goustin *et al*, 1986).

Cell growth, however, is not necessarily under control of positive regulating forces alone. For example, exogenous tumor growth factor-β (TGF-β) inhibits the growth of many cell types including neoplastic cells. Failure to express or respond to specific growth inhibitory factors that are released by cells to regulate orderly growth may also lead to disorders. There may be various possible causes of such a failure, including mutation or loss of the structural gene for TGF-β itself, loss of positive transcriptional or translational controls for expression of TGF-β or defects in the specific cellular receptor for

TGF-β. This alternative view, called "the Yin-Yang theory of cancer", inspired research on cloning tumor-suppressor genes to gain more information to support this theory (Massagué, 1985; Sporn and Roberts, 1985; Marx, 1986).

In some cases, growth factors have been shown to increase the transcription of proto-oncogenes, such as *myc* and *fos*, the products of which in their turn may regulate the transcription of other genes necessary for stimulation of cell proliferation (Cochran *et al*, 1984; Kruijer *et al*, 1984; Müller *et al*, 1984). These data suggest that many, if not all, of the oncogenic products may be involved in the pathway that follows the binding of the growth factor to its receptor. At each step of this pathway alterations may occur leading to the development of neoplastically transformed cells. A possible scenario of the sequence of events, that follow the binding of the growth factor to its receptor and lead to a cellular response can be outlined as follows (Figure 1):

* A growth factor binds to its cell surface receptor. After this step, the receptor may undergo allosteric changes, a redistribution in the membrane or an association with other membrane proteins. For example, the platelet-derived growth factor (PDGF) receptor consists of a PDGF binding domain, a transmembrane region and a cytoplasmic part that is able to bind adenosine triphosphate (ATP) and substrate(s) for phosphorylation (Frackelton *et al*, 1984; Daniel *et al*, 1985). In presence of PDGF, the receptor density on the cell surface decreases ("down-regulation") as the PDGF-receptor complex is internalized (Heldin *et al*, 1982; Nilsson *et al*, 1983).

  Activation of the receptor sometimes occurs in the ab-

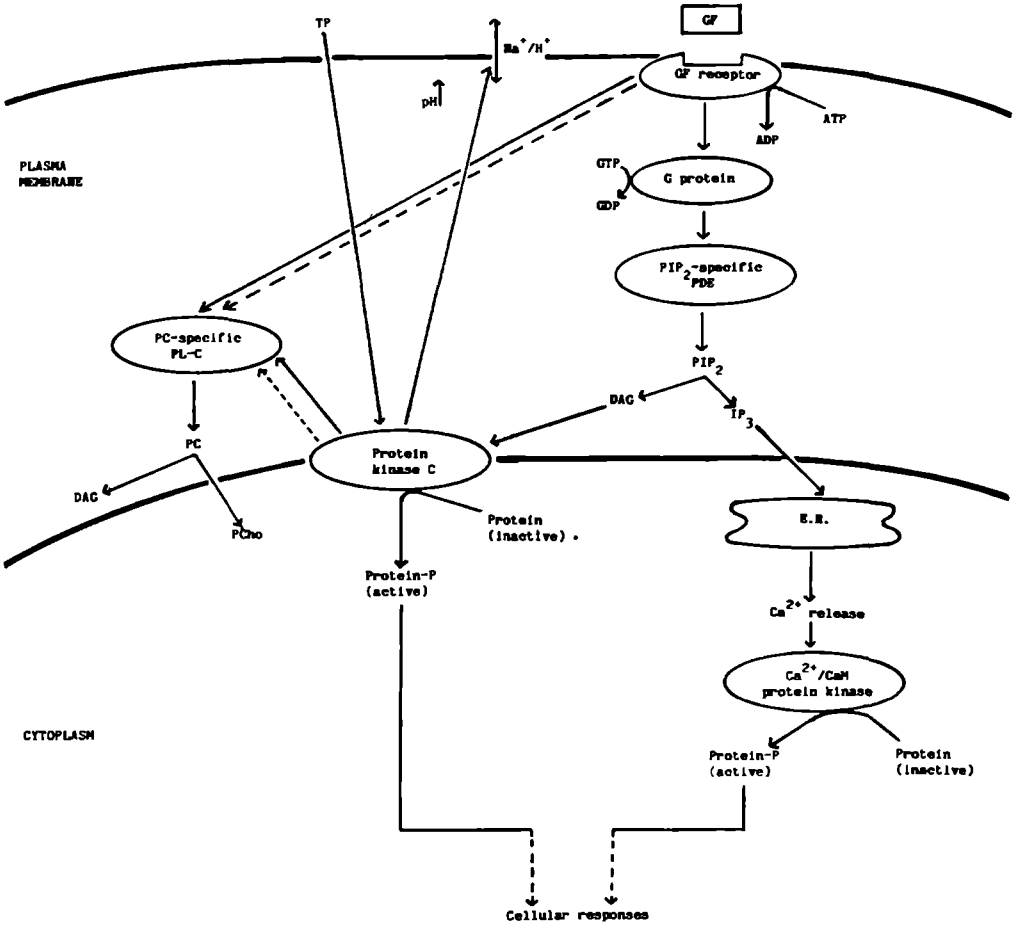


Figure 1. A possible mechanism of the influence of growth factors or tumor promoters on cellular responses. For a detailed description see text. Abbreviations: GF, growth factor; TP, tumor promoter; PIP_2-specific PDE, phosphatidylinositol-4,5-bisphosphate-specific phosphodiesterase; DAG, diacylglycerol; IP_3, inositol trisphosphate; E.R., endoplasmic reticulum; CaM, calmodulin; PC-specific PL-C, phosphatidylcholine-specific phospholipase C; PCho, phosphocholine; ATP, adenosine triphosphate; ADP, adenosine diphosphate; GTP, guanosine triphosphate; GDP, guanosine diphosphate.

sence of growth factor. Sequence homology between the viral oncogene v-erbB and the cellular receptor for EGF indicated that the main difference between the two is the absence of the EGF-binding domain in the viral oncogene product (Downward et al, 1984). Transformation by v-erB could therefore involve a mechanism in which the v-erbB protein relays a mitogenic signal even in the absence of growth factor binding.

* The activated growth factor receptor may transmit information through the plasma membrane into the cell by means of a family of G proteins: membrane proteins that are activated by binding guanosine triphosphate (GTP). Because of its location in the plasma membrane, its GTP-binding capacity and its involvement in the phosphatidylinositol-4,5-bisphosphate (PIP_2) breakdown pathway (see Figure 1 and below), the product of the ras proto-oncogene resembles a G protein (Fleischman et al, 1986).

The G protein activates an "amplifier" enzyme on the inner surface of the membrane. This enzyme converts precursor molecules into second messengers. In the cascade following binding of PDGF to its receptor, an "amplifier" enzyme is PIP_2-specific phosphodiesterase (PIP_2-specific PDE) which cleaves PIP_2 into the second messengers inositol trisphosphate (IP_3) and diacylglycerol (DAG) (Berridge, 1984a, 1985).

Addition of a tumor promoter (e.g. phorbol dibutyrate = PBT_2) to cells results in increased concentrations of DAG and phosphocholine (PCho), which is due to the activation by protein kinase C of phosphatidylcholine-specific phospholipase C (PC-specific PL-C). The observation that PDGF and

serum leads to the degradation of phosphatidylcholine into DAG and PCho, although the activation of protein kinase C is blocked, indicates that PDGF and serum act through a mechanism independent of protein kinase C. This is in contrast with PBT_2, which stimulates phosphatidylcholine hydrolysis in a manner solely dependent on protein kinase C (Besterman et al, 1986; Pasti et al, 1986).

* The second messengers activate cellular proteins by inducing conformational changes. One of the second messengers of the PIP_2 pathway is IP_3. Increased concentrations of IP_3 lead to a rapid release of calcium from an intracellular vesicular pool. The IP_3-sensitive calcium site of release is the endoplasmic reticulum and not the mitochondria as was concluded from cell fractionation and inhibition experiments. The release of calcium leads to phosphorylation of various proteins through the action of activated Ca^{2+}/CaM protein kinase (CaM: calmodulin) and, in this way, contributes to the cellular response on the binding of a growth factor to its receptor (Berridge, 1984a, 1985).

  The other second messenger of the PIP_2 system is the membrane bound DAG, which activates protein kinase C. Protein kinase C is capable of phosphorylating cytoplasmic proteins and causes alkalization of the cytoplasm by activation of a Na^+/H^+ exchange carrier, resulting into several physiological processes (Berridge, 1984a, 1985).

PDGF stimulation of quiescent cells is thought to follow one of the pathways described above leading to the phosphorylation of a large number of proteins as a consequence of protein

kinase C and Ca^{2+}/CaM protein kinase activation (Berridge et al, 1984b; Ek and Heldin, 1984). Binding of PDGF to its receptor causes also a rapid reorganization of the cytoskeleton actin and vinculin (Mellstrom et al, 1983; Burn et al, 1985; Herman and Pledger, 1985; Lassing and Lindberg, 1985). Evidence is provided for an interplay between PIP_2 and DAG with the polymerization and membrane interaction of actin and α-actinin respectively. It was shown that PIP_2 causes a rapid and efficient dissociation of profilactin with a concomitant polymerization of actin (Lassing and Lindberg, 1985). In vitro, a tight complex, displaying substructures similar to those of microfilament bundles in vivo, is formed between α-actinin, DAG and one fatty acid of a certain type (for example, palmitic acid) (Burn et al, 1985). These observations suggest a relationship between the increased activity in the PIP_2 cycle seen as a result of ligand-receptor interaction and the induction of actin filament formation.

Because the pathway of the signal transduction described above is very complex, it is conceivable that aberrations in this sequence of reactions can occur and might lead to abnormalities in cellular growth.

Platelet-derived growth factor (PDGF).

PDGF is contained in α-granules of platelets and is released during blood clotting or when platelets adhere at sites of blood vessel injury. PDGF may serve to promote wound healing since it is the most potent mitogen in serum for cells of mesenchymal origin, including fibroblasts, glial cells and smooth muscle cells. Because of its mitogen activity, PDGF may

also be involved in the formation of atherosclerotic lesions (Stiles, 1983; Deuel and Huang, 1984; Ross et al, 1986). Platelets are not the only cell type that release PDGF-like proteins; cultured rat aortic smooth muscle cells, human glioma cells, human sarcoma cells, human osteosarcoma cells, activated macrophages, cultured bovine endothelial cells, SSV-transformed cells (SSV: simian sarcoma virus) and activated human monocytes also produce such growth factors (Martinet et al, 1986; Ross et al, 1986).

Human PDGF is purified from platelets and biochemically and physiologically characterized. The major species are PDGF-I (31-35 kda) and PDGF-II (28-32 kda); they are equally active in bioassays, radioreceptor assays and radioimmuno assays but differ in their carbohydrate composition (Antoniades, 1981; Deuel et al, 1981; Huang et al, 1982). Reduction of PDGF results in loss of biological activity. The result of high-performance liquid chromatography experiments with reduced and alkylated PDGF suggested also that PDGF was a two-chain molecule. A heterogeneous PDGF-1 fraction (chain A; major form 18 kda) and a homogeneous PDGF-2 fraction (chain B; 16 kda) were found and the two chains appeared to be related to each other (Johnsson et al, 1982). These results led to the conclusion that PDGF might be a heterodimer consisting of one PDGF-1 and one PDGF-2 chain. However, Heldin et al (1986) recently isolated a PDGF-like growth factor resembling a homodimer of PDGF-1. This PDGF-like growth factor was isolated from a human osteosarcoma cell line and had structural, functional and immunological characteristics in common with PDGF isolated from human platelets (Heldin et al, 1980). The two chains of human

PDGF (PDGF-1 and -2) are encoded by genes which are located on different chromosomes (7 and 22 respectively) (Favera et al, 1982; Betsholtz et al, 1986). Screening of human tumor cell lines with specific probes for these two chains revealed that both genes are expressed independently. Immunoprecipitation of PDGF-like growth factors from conditioned medium of these cell lines correlates with the expression of PDGF-1 but not with that of PDGF-2, suggesting that all of the PDGF-like factors in these media are composed of only PDGF-1 chains (Betsholtz et al, 1986). Homodimers of both PDGF-1 and -2 chains are capable of binding to the PDGF receptor and exhibit mitogen activity (Heldin et al, 1980, 1986; Kelly et al, 1985). Therefore, the exact nature of the human PDGF subunit composition is still unknown.

PDGF receptor.

Several cell lines were shown to possess the capability of binding PDGF: human osteosarcoma cell lines, human fibroblasts, glial cells, smooth muscle cells and mouse 3T3 cells (Heldin et al, 1981, 1982; Bowen-Pope and Ross, 1982; Graves et al, 1984a, 1985; Williams et al, 1984).

The PDGF receptor is a 180 kda membrane glycoprotein, with an associated tyrosine-specific protein kinase activity. A less pronounced phosphorylation of a 130 kda protein was also noticed after binding of PDGF to its receptor. This 130 kda component is a proteolytic cleveage product of the 180 kda glycoprotein (Ek et al, 1982; Heldin et al, 1983; Frackelton et al, 1984; Daniel et al, 1985).

The genomic structure of the PDGF receptor has been charac-

terized using cDNA cloning (Yarden et al, 1986). From the deduced amino-acid sequence of the PDGF receptor it appeared that the structure of the receptor is closely related to the v-kit oncogene product and the receptor for the macrophage colony stimulating factor CSF-1, which is encoded by c-fms. Compared to the EGF receptor, an insertion is present in the kinase domain of v-kit, the PDGF and CSF-1 receptor. The cysteine residues in the extracellular domain of the PDGF receptor show a similar distribution pattern as those in the CSF-1 receptor (Yarden et al, 1986).

Binding of PDGF to its receptor leads to a rapid internalization of the factor via small vesicles. Because the internalization is followed by a tempory loss of the ability of the cells to bind the ligand ("down-regulation"), the latter is probably ingested as a receptor-ligand complex (Pastan and Willingham, 1981; Huang et al, 1982; Nilsson et al, 1983). After PDGF binding, several "immediate" and "early" events can be monitored. Some "immediate" events, which are independent of RNA synthesis, are: tyrosine-specific phosphorylation, stimulation of phosphatidylinositol turnover and reorganization of actin filaments (Ek et al, 1982; Berridge, 1984b; Herman and Pledger, 1985). Examples of "early", transcriptionally dependent events are: induction of transcription of genes among which c-fos and c-myc, presence of unique proteins and division of cells without further requirement of PDGF (Rutherford and Ross, 1976; Pledger et al, 1981; Olashow and Pledger, 1983; Cochran et al, 1984; Kruijer et al, 1984; Müller et al, 1984).

Because of the homology of the amino-acid sequence of PDGF-2 with the predicted amino-acid sequence of the transforming pro-

tein encoded by the v-sis oncogene of SSV and the inhibition of acute transformation by SSV using antibodies against PDGF (Favera et al, 1981; Wong-Staal et al, 1981; Robbins et al, 1982; Doolittle et al, 1983; Waterfield et al, 1983; Johnsson et al, 1985a), it was of interest to define further the c-sis proto-oncogene and the sequences controlling its expression.

The sis proto-oncogene.

The human c-sis proto-oncogene is located on chromosome 22 (q11,qter) (Goyns, 1983; Groffen et al, 1983). Human chromosomes 22 and 9 are implicated in the chromosome translocation (t(9,22)(q34;q11)), observed in an Ewing-Sarcoma cell line and in patients suffering from chronic myeloid leukemia (CML) (Goyns, 1983; Groffen et al, 1983; Geurts Van Kessel, 1985). The sis proto-oncogene is translocated to chromosome 9 and the abl proto-oncogene to chromosome 22 leading to the Philadelphia chromosome (Groffen et al, 1984). The results that no sis mRNA transcripts but, instead, altered c-abl transcripts fused with the bcr gene were observed led to the conclusion that the c-sis proto-oncogene was not involved in the formation of these tumors (Bechet et al, 1984; Heisterkamp et al, 1985; Shtivelman et al, 1985).

Transcription of the sis proto-oncogene is observed in several cell types: endothelial cells, HeLa cells, osteosarcoma cell line, prostatic carcinoma cell line, glioblastoma cell line, activated monocytes, lung tumor cell line, placenta, neuroblastoma cell line (Barrett et al, 1984; Graves et al, 1984b; Goustin et al, 1985; Van den Ouweland et al, 1985, 1986, 1987; Van Zoelen et al, 1985; Martinet et al, 1986). As the main

transcriptional product in human cells, an mRNA species of about 3.5 kb with a coding potential for a protein of about 27 kda , the sequence of which shows strong homology with PDGF-2, was described (Collins et al, 1985; Ratner et al, 1985; Rao et al, 1986). Because of the striking difference between the size of the v-sis oncogene (1 kbp) and the mRNA transcribed of the c-sis proto-oncogene (3.5 kb) it is clear that the complexity of the c-sis proto-oncogene is greater than the v-sis homologous sequences (Devare et al, 1983; Van den Ouweland et al, 1986).

As mentioned above, the c-sis proto-oncogene may be involved in differentiation and transformation. Jay et al (1985) have shown that during endothelial cell differentiation in vitro expression of the sis gene was modulated. Transformed cells obtained upon transfection with cDNA or cloned genomic DNA of the sis locus were morphologically identical to those observed upon transfection with the viral oncogene v-sis, suggesting that no changes such as point mutations were required for transformation (Clarke et al, 1984; Gazit et al, 1984). Antibodies against PDGF were used in an attempt to identify the transforming product in the transfection assays with SSV. The presence of anti-PDGF antibodies appeared to inhibit both proliferation and SSV-induced morphological changes in human fibroblasts, suggesting that SSV-transformation is due solely to the autocrine action of a PDGF agonist. Despite a continuous expression of the viral genome and production of a PDGF-like factor leading to a "down-regulation" of the PDGF receptor the SSV-transformed cells at high passage level revert phenotypically. These data give reason to believe that SSV-transformed cells

and their normal counterparts senesce by the same mechanism involving a block in the post-receptor pathway. Taken together, these data indicate that SSV-transformation must be mediated by a growth factor that mimics PDGF (Huang *et al*, 1984; Johnsson *et al*, 1985a, 1985b, 1986).

This thesis describes the moleculary cloning and characterization of the human and feline c-*sis* proto-oncogenes and their flanking sequences. To gain more insight in the regions that have regulatory functions involved in the expression of the c-*sis* proto-oncogene, a good characterization of this gene is necessary. As already described, the simian sarcoma virus has captured only limited sequences from its cellular counterpart and, therefore, the viral oncogene can not be used to define the entire c-*sis* proto-oncogene. Taken into account the conserved nature of the proto-oncogenes (R. Weiss *et al*, 1984), comparative analysis of the c-*sis* proto-oncogenes of two species could lead to a complete description of the gene. The species used in this thesis are human and cat. The latter was chosen because of the existense of a v-*sis* homologous sequence present in the Parodi-Irgens strain and the great knowledge of genetic maps of this species (Besmer *et al*, 1983; O'Brien *et al*, 1985).

REFERENCES

- Antoniades, H.N. (1981) Proc. Natl. Acad. Sci. USA 78, 7314-7317.
- Barnes, D. and Sato, G. (1980) Cell 22, 649-655.
- Barrett, T.B., Gajdusek, C.M., Schwartz, S.M., McDougall, J.K. and Benditt, E.P. (1984) Proc. Natl. Acad. Sci. USA 81, 6772-6774.
- Bechet, J-M., Bornkamm, G., Freese, U-K and Lenoir, G.M. (1984) New England J. of Medicine 310, 393.
- Berridge, M.J. (1984a) Biochem. J. 220, 345-360.
- Berridge, M.J., Heslop, J.P., Irvine, R.F. and Brown, K.D. (1984b) Biochem. J. 222, 195-201.
- Berridge, M.J. (1985) Sci. Am. 253, 124-134.
- Besmer, P., Snyder, H.W.Jr., Murphy, J.E., Hardy, W.D.Jr. and Parodi, A. (1983) J. Virology 46, 606-613.
- Besterman, J.M., Duronio, V. and Cuatrecasas, P. (1986) Proc. Natl. Acad. Sci. USA 83, 6785-6789.
- Betsholtz, C., Johnsson, A., Heldin, C-H., Westermark, B., Lind, P., Urdea, M.S., Eddy, R., Shows, T.B., Philpott, K., Mellor, A.L., Knott, T.J. and Scott, J. (1986) Nature 320, 695-699.
- Bishop, J.M. (1983) Ann. Rev. Biochem. 52, 301-354.
- Bowen-Pope, D.F. and Ross, R. (1982) J. Biol. Chem. 257, 5161-5171.
- Burn, P., Rotman, A., Meyer, R.K. and Burger, M.M. (1985) Nature 314, 469-472.
- Clarke, M.F., Westin, E., Schmidt, D., Josephs, S.F., Ratner, L., Wong-Staal, F., Gallo, R.C. and Reitz, M.S.Jr. (1984) Nature 308, 464-467.
- Cochran, B.H., Zullo, J., Verma, I.M. and Stiles, C.D. (1984) Science 226, 1080-1082.
- Collins, T., Ginsburg, D., Boss, J.M., Orkin, S.H. and Pober, J.S. (1985) Nature 316, 748-750.
- Daniel, T.O., Tremble, P.M., Frackelton, A.R.Jr. and Williams, L.T. (1985) Proc. Natl. Acad. Sci. USA 82, 2684-2687.
- Deuel, T.F., Huang, J.S., Proffitt, R.T., Baenziger, J.U., Chang, D. and Kennedy, B.B. (1981) J. Biol. Chem. 256, 8896-8899.
- Deuel, T.F., Huang, J.S., Huang, S.S., Stroobant, P. and Waterfield, M.D. (1983) Science 221, 1348-1350.
- Deuel, T.F. and Huang. J.S. (1984) J. Clin. Invest. 74, 669-676.

- Devare, S.G., Reddy, E.P., Law, J.D., Robbins, K.C. and Aaronson, S.A. (1983) Proc. Natl. Acad. Sci. USA 80, 731-735.
- Doolittle, R.F., Hunkapiller, M.W., Hood, L.E., Devare, S.G., Robbins, K.C., Aaronson, S.A. and Antoniades, H.N. (1983) Science 221, 275-277.
- Downward, J., Yarden, Y., Mayes, E., Scrace, G., Totty, N., Stockwell, P., Ullrich, A., Schlessinger, J. and Waterfield, M.D. (1984) Nature 307, 521-527.
- Ek, B., Westermark, B., Wasteson, A. and Heldin, C-H. (1982) Nature 295, 419-420.
- Ek, B. and Heldin, C-H. (1984) J. Biol. Chem. 259, 11145-11152.
- Favera, R.D., Gelmann, E.P., Gallo, R.C. and Wong-Staal, F. (1981) Nature 292, 31-35.
- Favera, R.D., Gallo, R.C., Giallongo, A. and Croce, C.M. (1982) Science 218, 686-688.
- Fishinger, P.J. (1982) in Molecular Biology of RNA Tumor Viruses (ed. Stephenson, J.R.) Academic Press, NY, 162-250.
- Fleischman, L.F., Chahwala, S.B. and Cantley, L. (1986) Science 231, 407-410.
- Frackelton, A.R.Jr., Tremble, P.M. and Williams, L.T. (1984) J. Biol. Chem. 259, 7909-7915.
- Gazit, A., Igarashi, H., Chiu, I-M., Srinivasan, A., Yaniv, A., Tronick, S.R., Robbins, K.C. and Aaronson, S.A. (1984) Cell 39, 89-97.
- Geurts Van Kessel, A., Turc-Carel, C., De Klein; A., Grosveld, G., Lenoir, G. and Bootsma, D. (1985) Mol. Cell. Biol. 5, 427-429.
- Goustin, A.S., Betsholtz, C., Pfeifer-Ohlsson, S., Persson, H., Rydnert, J., Bywater, M., Holmgren, G., Heldin, C-H., Westermark, B. and Ohlsson, R. (1985) Cell 41, 301-312.
- Goustin, A.S., Leof, E.B., Shipley, G.D. and Moses, H.L. (1986) Cancer Res. 46, 1015-1029.
- Goyns, M.H. (1983) Bioscience Reports 3, 1107-1111.
- Graves, D.T., Antoniades, H.N., Williams, S.R. and Owen, A.J. (1984a) Cancer Res. 44, 2966-2970.
- Graves, D.T., Owen, A.J., Barth, R.K., Tempst, P., Winoto, A., Fors, L. and Hood, L.E. (1984b) Science 226, 972-974.
- Graves, D.T., Owen, A.J. and Antoniades, H.N. (1985) Biochem. Biophys. Res. Comm. 129, 56-62.
- Greene, W.C., Leonard, W.J., Wano, Y., Svetlik, P.B., Peffer, N.J., Sodroski, J.G., Rosen, C.A., Goh, W.C. and

Haseltine, W.A. (1986) Science 232, 877-880.
- Groffen, J., Heisterkamp, N., Stephenson, J.R., Geurts Van Kessel, A., De Klein, A., Grosveld, G. and Bootsma, D. (1983) J. Exp. Medicine 158, 9-15.
- Groffen, J., Stephenson, J.R., Heisterkamp, N., De Klein, A., Bartram, C.R. and Grosveld, G. (1984) Cell 36, 93-99.
- Heisterkamp, N., Stam, K., Groffen, J., De Klein, A. and Grosveld, G. (1985) Nature 315, 758-761.
- Heldin, C-H., Westermark, B. and Wasteson, A. (1980) J. Cell. Physiol. 105, 235-246.
- Heldin, C-H., Westermark, B. and Wasteson, A. (1981) Proc. Natl. Acad. Sci. USA 78, 3664-3668.
- Heldin, C-H., Wasteson, A. and Westermark, B. (1982) J. Biol. Chem. 257, 4216-4221.
- Heldin, C-H., Ek, B. and Ronnstrand, L. (1983) J. Biol. Chem. 258, 10054-10061.
- Heldin, C-H. and Westermark, B. (1984) Cell 37, 9-20.
- Heldin, C-H., Johnsson, A., Wennergren, S., Wernstedt, C., Betsholtz, C. and Westermark, B. (1986) Nature 319, 511-514.
- Herman, B. and Pledger, W.J. (1985) J. Cell Biol. 100, 1031-1040.
- Huang, J.S., Huang, S.S., Kennedy, B. and Deuel, T.F. (1982) J. Biol. Chem. 257, 8130-8136.
- Huang, J.S., Huang, S.S. and Deuel, T.F. (1984) Cell 39, 79-87.
- Hunter, T. (1984) Sci. American 251, 60-69.
- Jaye, M., McConathy, E., Drohan, W., Tong, B., Deuel, T. and Maciag, T. (1985) Science 228, 882-885.
- Johnsson, A., Heldin, C-H., Westermark, B. and Wasteson, A. (1982) Biochem. Biophys. Res. Comm. 104, 66-74.
- Johnsson, A., Betsholtz, C., Heldin, C-H. and Westermark, B. (1985a) Nature 317, 438-440.
- Johnsson, A., Betsholtz, C., von der Helm, K., Heldin, C-H. and Westermark, B. (1985b) Proc. Natl. Acad. Sci. USA 82, 1721-1725.
- Johnsson, A., Betsholtz, C., Heldin, C-H. and Westermark, B. (1986) EMBO J. 5, 1535-1541.
- Kelly, J.D., Raines, E.W., Ross, R. and Murray, M.J. (1985) EMBO J. 4, 3399-3405.
- Kruijer, W., Cooper, J.A., Hunter, T. and Verma, I.M. (1984) Nature 312, 711-716.
- Lassing, I and Lindberg, U. (1985) Nature 314, 472-474.
- Martinet, Y., Bitterman, P.B., Mornex, J-F., Grotendorst,

G.R., Martin, G.R. and Crystal, R.G. (1986) Nature 319, 158-160.
- Marx, J.L. (1986) Science 232, 1093-1095.
- Massagué, J. (1985) TIBS 10, 237-240.
- Mellstrom, K., Hoglund, A-S., Nistér, M., Heldin, C-H., Westermark, B. and Lindberg, U. (1983) J. Muscle Res. and Cell Motility 4, 589-609.
- Müller, R., Bravo, R., Burckhardt, J. and Curran, T. (1984) Nature 312, 716-720.
- Nilsson, J., Thyberg, J., Heldin, C-H., Westermark, B. and Wasteson, A. (1983) Proc. Natl. Acad. Sci. USA 80, 5592-5596.
- Nilsson, J., Sjolund, M., Palmberg, L., Thyberg, J. and Heldin, C-H. (1985) Proc. Natl. Acad. Sci. USA 82, 4418-4422.
- O'Brien, S.J., Seuánez, H.N. and Womack, J.E. (1985) in Molecular Evolutionary Genetics (Ed. MacIntyre, R.J.) Plenum Publ. Corp. 519-589.
- Olashaw, N.E. and Pledger, W.J. (1983) Nature 306, 272-274.
- Pastan, I.H. and Willingham, M.C. (1981) Science 214, 504-509.
- Pasti, G., Lacal, J-C., Warren, B.S., Aaronson, S.A. and Blumberg, P.M. (1986) Nature 324, 375-377.
- Pledger, W.J., Hart, C.A., Locatell, K.L. and Scher, C.D. (1981) Proc. Natl. Acad. Sci. USA 78, 4358-4362.
- Rall, L.B., Scott, J., Bell, G.I., Crawford, R.J., Penschow, J.D., Niall, H.D. and Coghlan, J.P. (1985) Nature 313, 228-231.
- Rao, C.D., Igarashi, H., Chiu, I-M., Robbins, K.C. and Aaronson, S.A. (1986) Proc. Natl. Acad. Sci. USA 83, 2392-2396.
- Ratner, L., Josephs, S.F., Jarrett, R., Reitz, M.S.Jr. and Wong-Staal, F. (1985) Nucl. Acids Res. 13, 5007-5018.
- Robbins, K.C., Hill, R.L. and Aaronson, S.A. (1982) J. Virology 41, 721-725.
- Ross, R., Raines, E.W. and Bowen-Pope, D.F. (1986) Cell 46, 155-169.
- Rutherford, R.B. and Ross, R. (1976) J. Cell Biol. 69, 196-203.
- Sap, J., Munoz, A., Damm, K., Goldberg, Y., Ghysdael, J., Leutz, A., Beug, H. and Vennstrom, B. (1986) Nature 324, 635-640.
- Sherr, C.J., Rettenmier, C.W., Sacca, R., Roussel, M.F., Look, A.T. and Stanley, E.R. (1985) Cell 41, 665-676.

- Shtivelman, E., Lifshitz, B., Gale, R.P. and Canaani, E. (1985) Nature 315, 550-554.
- Sporn, M.B. and Roberts, A.B. (1985) Nature 313, 745-747.
- Stiles, C.D. (1983) Cell 33, 653-655.
- Tsao, M.C., Walthall, B.J. and Ham, R.G. (1982) J. Cell. Physiol. 110, 219-229.
- Van den Ouweland, A.M.W., Breuer, M.L., Steenbergh, P.H., Schalken, J.A., Bloemers, H.P.J. and Van de Ven, W.J.M. (1985) Biochim. Biophys. Acta 825, 140-147.
- Van den Ouweland, A.M.W., Roebroek, A.J.M., Schalken, J.A., Claesen, C.A.A., Bloemers, H.P.J. and Van de Ven, W.J.M. (1986) Nucl. Acids Res. 14, 765-778.
- Van den Ouweland, A.M.W., Van Groningen, J.J.M., Schalken, J.A., Bloemers, H.P.J. and Van de Ven, W.J.M. (1987) Nucl. Acids Res. 15, 959-970.
- Van Zoelen, E.J.J., Van de Ven, W.J.M., Franssen, H.J., Van Oostwaard, Th.M.J., Van der Saag, P.T., Heldin, C-H. and de Laat, S.W. (1985) Mol. Cell. Biol. 5, 2289-2297.
- Walthall, B.J. and Ham, R.G. (1981) Exp. Cell Res. 134, 303-311.
- Waterfield, M.D., Scrace, G.T., Whittle, N., Stroobant, P., Johnsson, A., Wasteson, A., Westermark, B., Heldin, C-H., Huang, J.S. and Deuel, T.F. (1983) Nature 304, 35-39.
- Weinberg, R.A. (1985) Science 230, 770-776.
- Weinberger, C., Thompson, C.C., Ong, E.S., Lebo, R., Gruol, D.J. and Evans, R.M. (1986) Nature 324, 641-646.
- Weiss, R., Teich, N., Varmus, H. and Coffin, J. (1984) in RNA Tumor Viruses, Molecular Biology of Tumor Viruses, Part 1 and 2, Cold Spring Harbor Laboratory, U.S.A.
- Williams, L.T., Tremble, P.M., Lavin, M.F. and Sunday, M.E. (1984) J. Biol. Chem. 259, 5287-5294.
- Wong-Staal, F., Favera, R.D., Gelmann, E.P., Manzari, V., Szala, S., Josephs, S.F. and Gallo, R.C. (1981) Nature 294, 273-275.
- Wong-Staal, F. and Gallo, R.C. (1985) Nature 317, 395-403.
- Yarden, Y., Escobedo, J.A., Kuang, W-J., Yang-Feng, T.L., Daniel, T.O., Tremble, P.M., Chen, E.Y., Ando, M.E., Harkins, R.N., Francke, U., Fried, V.A., Ullrich, A. and Williams, L.T. (1986) Nature 323, 226-232.

CHAPTER 2

COMPARATIVE ANALYSIS OF THE HUMAN AND FELINE c-*sis* PROTO-ONCOGENES. IDENTIFICATION OF 5' HUMAN c-*sis* CODING SEQUENSES THAT ARE NOT HOMOLOGOUS TO THE TRANSFORMING GENE OF SIMIAN SARCOMA VIRUS.

Biochimica et Biophysica Acta 825 (1985), 140-147.

Biochimica et Biophysica Acta 825 (1985) 140–147
Elsevier

BBA 91458

Comparative analysis of the human and feline c-*sis* proto-oncogenes. Identification of 5′ human c-*sis* coding sequences that are not homologous to the transforming gene of simian sarcoma virus

Ans M.W. van den Ouweland [a], Marco L. Breuer [a], Paul H. Steenbergh [b], Jack A. Schalken [a], Henri P.J. Bloemers [a] and Wim J.M. Van de Ven [a] *

[a] *Molecular Oncology Section, Department of Biochemistry, University of Nijmegen, Nijmegen, and* [b] *Institute of Molecular Biology, University of Utrecht, Utrecht (The Netherlands)*

(Received January 3rd, 1985)

Key words Proto-oncogene Restriction enzyme analysis, (Feline Human)

Feline and human genetic sequences, homologous to the v-*sis* gene of simian sarcoma virus, have been isolated from cosmid gene libraries and characterized by restriction endonuclease analysis. Comparison of the two loci revealed their related structural organization. In both loci, similar unique genetic sequences were found upstream of the v-*sis* homologous region and these hybridized to a 4.2 kb c-*sis* transcript in human lung tumor cells. These data establish and map as yet unidentified coding sequences at the 5′ part of the c-*sis* proto-oncogene of both species.

Introduction

It is well established that viral oncogenes of acutely transforming RNA tumor viruses have arisen by recombination between retroviral genomes and genetic sequences of cellular origin [1–3] These cellular sequences, which are highly conserved during evolution, are known as proto-oncogenes In their normal form, they are thought to have an important function during growth and differentiation [4–7] It has also been suggested that proto-oncogenes are involved in tumorigenesis [8–11] In that case, however, genetic changes such as point mutations [12], translocations [13] or insertions [14] seem to be required Some viral oncogenes code for proteins that have been identified and characterized as a protein kinase capable of phosphorylating tyrosine residues [15,16], a growth factor [17,18] or a growth factor receptor [19]

The proto-oncogene under investigation in the present study is c-*sis* Genetic sequences from this locus have been captured in the genome of Parodi-Irgens feline sarcoma virus (PI-FeSV), isolated from a fibrosarcoma of a cat infected with feline leukemia virus (FeLV) [20], and simian sarcoma virus (SSV), isolated from a fibrosarcoma of a woolly monkey [21] Their transforming proteins have been identified as a 76 kDa fusion protein containing *gag* and *sis* genetic sequences [22] and 28 kDa protein [23–25], designated $p28^{sis}$, respectively Recently, comparison of amino acid sequences of human PDGF-2 with predicted sequences of $p28^{sis}$ has shown that these proteins share a virtually identical region [17,18,25] In the light of the fact that in certain human tumors transcripts have been found that are related to v-*sis* [26,27] and that a number of human tumors

* To whom correspondence should be addressed

Abbreviations RFLP, restriction fragment length polymorphism ALL, acute lymphocytic leukemia, SSV simian sarcoma virus PI FeSV Parodi-Irgens feline sarcoma virus FeLV feline leukemia virus

produce PDGF-like proteins [28,29], it is of interest to investigate the potential involvement of the *sis* locus in tumorigenesis As a first approach, studies were focussed upon the isolation and characterization of this locus and its flanking cellular sequences The size of the viral oncogene in SSV is about 1 kbp [23,24,30] and homologous transcripts found in human tumor cell lines are about 4 2 kb and 2 7 kb in size [27] This could indicate that the human c-*sis* locus contains more genetic sequences than those homologous to v-*sis* To resolve this, we have molecularly cloned the human and feline c-*sis* proto-oncogenes for comparative analysis and have performed Northern blot analysis of c-*sis* transcripts in human tumor cells

Materials and Methods

Cosmid, phages, plasmids and E coli strains

pC60, a recombinant of pBR322 containing the entire SSV provirus [31], was a gift from R C Gallo FAO1 and FAO2 represent subclones of pC60 in phage vector m13mp8 They contain the 0 5 kbp *Sac*I/*Sma*I or the 0 5 kbp *Sma*I/*Xba*I restriction fragments of pC60, respectively Both fragments together represent almost the complete v-*sis* oncogene Restriction endonuclease *Sac*I cleaves the oncogene at nucleotide position 22 and *Xba*I just before the last nucleotide of v-*sis* *Bam*HI linkers were used in preparation of the subclones

pAO70, pAO73, pAO79 and pAO121 are subclones of the human c-*sis* locus (see also Fig 3) pAO70 consists of a 5 0 kbp *Hin*dIII DNA fragment subcloned in pSVBR94 An internal 3 8 kbp *Kpn*I/*Kpn*I fragment of this insert was used as a molecular probe pAO73 and pAO79 are recombinant clones of pBR322 containing respectively an 1 7 kbp *Bam*HI and an 1 3 kbp *Bam*HI/*Hin*dIII restriction fragment A 2 0 kbp *Eco*RI/*Hin*dIII restriction fragment was subcloned in pUC18 resulting in pAO121

E coli strain HB101 (hsdR$^-$, recA$^-$) was used as a host for pC60, pAO70, pAO73, pAO79 and pAO121 *E coli* strain 1046 (supE, supF, hsdS$^-$, recA$^-$) was used for preparing the pJB8 [32] cosmid libraries The recombinant m13mp8 bacteriophages FAO1 and FAO2 were propagated in *E coli* strain JM101 [33]

Preparation of DNA probes and hybridization

Preparation of DNA probes and their nick translation [34] was carried out as described before [35] Primer extension reactions were according to Messing and co-workers [33] The specific activity of the probes used in the hybridization studies was (2–5) 10^8 cpm/μg Agarose gel electrophoresis, Southern blot [36] and hybridization analysis was performed as described before [35]

Molecular cloning of c-sis sequences

The human and feline v-*sis* homologous genetic sequences were isolated from previously constructed cosmid gene libraries [35,37] High-molecular-weight DNA used for the construction of the cosmid gene libraries was prepared from blood cells of a patient suffering from a common type acute lymphocytic leukemia (ALL, patient No 1283) and from lung tissue of a domestic cat Southern blot analysis of DNA isolated from the leukemia patient indicated that the c-*sis* proto-oncogene of this patient was similar in structure to that found in normal human liver DNA (data not shown) and, thus, the previously constructed library was suitable for the isolation of the c-*sis* proto-oncogene

RNA isolation and Northern blotting

Total cellular RNA was isolated according to the guanidine-thiocyanate procedure described by Feramisco et al [38] Upon poly(A) selection by oligo(dT)-cellulose chromatography, RNA was dissolved in sodium phosphate buffer (25 mM, pH 6 5), which contained DMSO (50%) and glyoxal (0 5 M), and heated to 50°C for 1 h RNA was size fractionated by agarose gel electrophoresis (1 4% agarose gel) and transferred to nitrocellulose for hybridization analysis as described before [39]

Results

Analysis of genetic sequences homologous to v-sis in human and feline genomic DNA

For analysis of the size and distribution of genetic sequences in human and feline genomic DNA homologous to v-*sis* two v-*sis*-specific probes, FAO1 and FAO2 were used These probes contained the 5′ and 3′ half of the viral oncogene as described under Materials and Methods Hu-

man DNA, isolated from the ALL patient (No. 1283) and digested with restriction endonuclease *Hin*dIII (Fig. 1, lane 1), revealed a hybridizing fragment of about 23 kbp. A similar analysis of DNA isolated from blood cells of a normal individual (No. 040883) revealed three hybridizing fragments of about 23 kbp, 17 kbp and 5.7 kbp (Fig. 1, lane 2), indicating the presence of a *Hin*dIII restriction fragment length polymorphism (RFLP). Among 10 human DNA preparations tested in this way, 6 had the same *Hin*dIII RFLP in one allele, but none of the DNAs showed the *Hin*dIII RFLP in both alleles (data not shown). Southern blot analysis of DNA of ALL patient No. 1283 with restriction endonuclease *Eco*RI revealed a single hybridizing fragment of about 25 kbp, and with restriction endonuclease *Xba*I, three v-*sis* homologous DNA fragments of about 25 kbp, 8.4 kbp and 3.6 kbp were identified (data not shown).

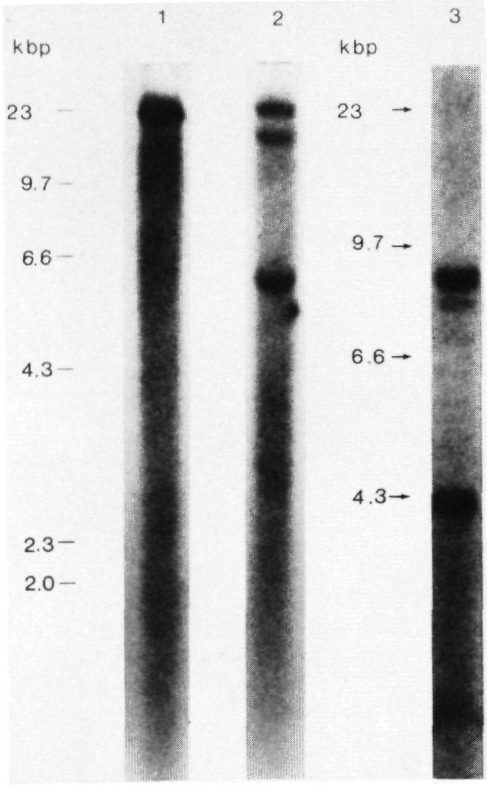


Fig. 1. Identification of genetic sequences in human and feline cellular DNA homologous to v-*sis*. High-molecular-weight DNA was prepared from blood cells of a patient suffering from a common type acute lymphocytic leukemia (patient No. 1283) (lane 1) and blood cells of a normal individual (lane 2) or from lung tissue of a domestic cat (cat No. 031182) (lane 3), digested with restriction endonuclease *Hin*dIII (lane 1 and 2) or *Eco*RI (lane 3) and size-fractionated by agarose gel electrophoresis through an 0.7% agarose gel. Upon transfer of the DNA to nitrocellulose, hybridization analysis was performed with ^{32}P-labeled probes (FAO1 and FAO2) representing almost the complete v-*sis* oncogene. Molecular weight markers included are *Hin*dIII-digested λ DNA fragments.

Southern blot analysis of high-molecular-weight feline DNA (cat No. 031182) with FAO1 and FAO2 is shown in lane 3 of Fig. 1. Two *Eco*RI hybridizing fragments could be identified. One was about 8 kbp in size and the other 4.3 kbp. Southern blot analysis with restriction endonuclease *Bam*HI revealed a 13 kbp, a 4.8 kbp and an 1.3 kbp v-*sis* homologous DNA fragment (data not shown). Similarly, DNA fragments of about 25 kbp, 5 kbp and 2.6 kbp were found using restriction endonuclease *Hin*dIII (data not shown).

These results indicate that in both species the v-*sis* homologous genetic sequences are distributed discontinuously over a relatively large region. Similar observations have previously been reported for the human locus [40–42]. Furthermore, in view of the fact that the viral oncogene of SSV seems to represent only a minor portion of the human cellular locus, the size of the complete proto-oncogene could be much greater.

Molecular cloning of human and feline c-sis genetic sequences

Based upon the results described above, we have used a previously described human and feline cosmid gene library to isolate from both species a large contiguous DNA region that contained the c-*sis* proto-oncogenes. Screening of the human cosmid gene library (300 000 colonies) with the v-*sis*-specific probes FAO1 and FAO2 resulted in one c-*sis* containing cosmid clone (ALLW-1283-C121) as is shown in Fig. 2A, lane 1 and Fig. 2B, lane 1. ALLW-1283-C121 contained all human v-*sis* homologous sequences as detected in genomic blot analysis (hybridization data not shown). A restriction map of this clone was obtained on the basis of

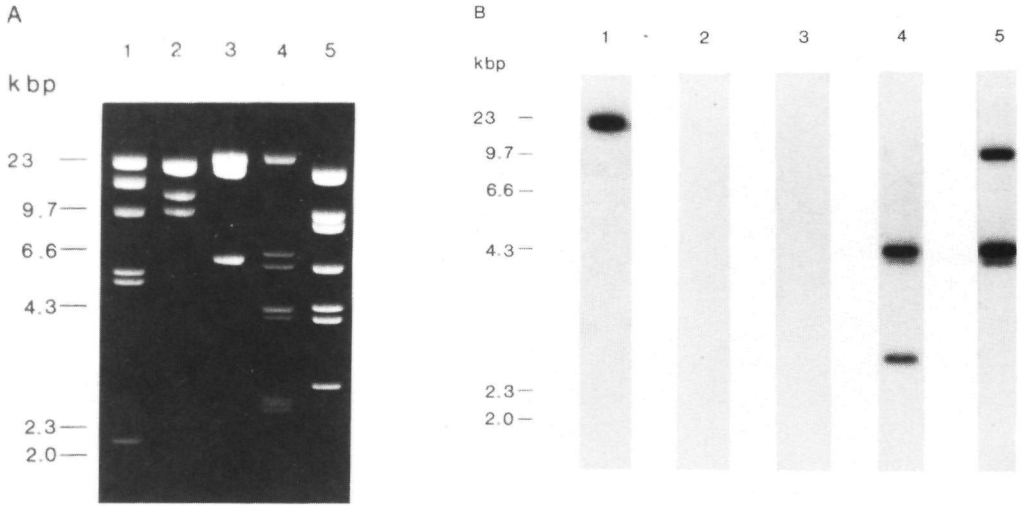


Fig. 2. Characterization of cosmid clones isolated from a human and feline cosmid gene library. (A) DNA of cosmid clones ALLW-1283-Cl21 (lane 1), MB10 (lane 2), MB25 (lane 3) and feline cosmid clones MB65 (lane 4) and MB70 (lane 5) was digested with restriction endonuclease *Hin*dIII (lanes 1, 2 and 3) or with *Eco*RI (lanes 4 and 5), size-fractionated by agarose gel electrophoresis (0.6% agarose gel) and visualized by ethidium bromide staining. (B) Southern blot analysis of human cosmid clones ALLW-1283-Cl21 (lane 1), MB10 (lane 2), MB25 (lane 3) and feline cosmid clones MB65 (lane 4) and MB70 (lane 5) with v-*sis*-specific molecular probes FAO1 and FAO2. Digestion with restriction endonucleases is as described under A. Cosmid clones MB10 and MB25 were isolated by 'chromosome walking' (see Fig. 3) using the insert of pAO79 as a probe. Molecular weight markers were as described in the legend to Fig. 1.

restriction endonuclease digestion patterns using the restriction endonucleases *Bam*HI, *Eco*RI, *Hin*dIII, *Kpn*I, *Xba*I and *Xho*I or combinations of these in sequential digestions (data summarized in Fig. 3). The orientation of c-*sis* sequences of ALLW-1283-Cl21 was obtained by hybridization analysis using FAO1 or FAO2 as a 5′ or 3′ v-*sis*-specific probe (data summarized in Fig. 3).

For the purpose of 'chromosome walking', a unique 5′-specific (pAO79) and 3′-specific (pAO73) probe was prepared from clone ALLW-1283-Cl21 and used in further screening of the human cosmid library. Two positive clones, designated MB10 and MB25, were found (Fig. 2A, lanes 2 and 3, and Fig. 2B, lanes 2 and 3), both extending in a direction upstream of clone ALLW-1283-CI21 (Fig. 3). The inserts of these cosmid clones did not hybridize with v-*sis* (Fig. 2B, lanes 2 and 3). No positive cosmid clones were found after screening another 200 000 colonies with pAO73.

In parallel experiments, two cosmid clones containing c-*sis* sequences were obtained upon screening of the feline cosmid library (500 000 colonies) with probes FAO1 and FAO2 (Fig. 2A, lanes 4 and 5; Fig. 2B, lanes 4 and 5). The complete feline v-*sis* cellular homolog was present in these cosmid clones. Determination of the orientation of feline c-*sis* and construction of its restriction enzyme map was done in the same way as described for the human c-*sis* clones (Fig. 3).

As summarized in Fig. 3, the three human cosmid clones contain overlapping sequences corresponding to a single contiguous region of human cellular DNA of about 65 kbp. The distribution of the v-*sis* homologous region is in accordance with previous reports [40–42] in which a total of six exons were identified. The feline DNA inserts represent about 60 kbp of contiguous cellular DNA and the organization of the v-*sis* homologous regions seems similar to those in the human locus.

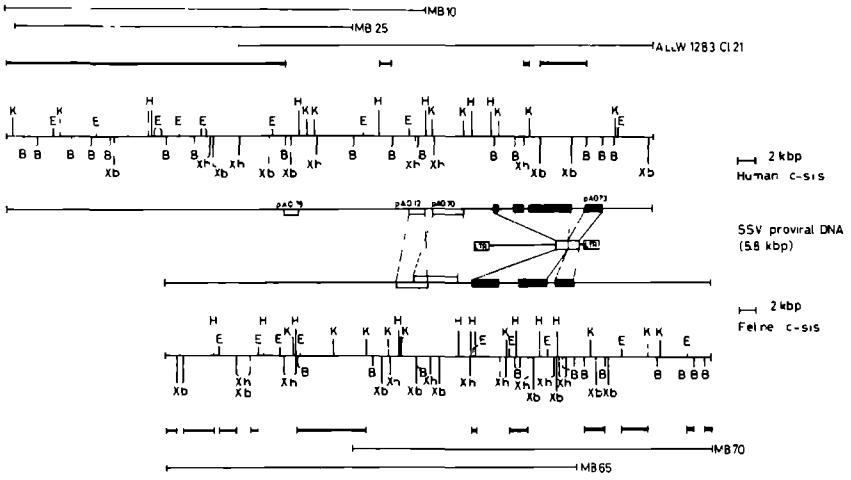


Fig. 3. Restriction endonuclease map of the human and feline c-*sis*-containing DNA regions In the upper part of the figure, cellular DNA inserts of the human cosmid clones ALLW-1283-Cl21, MB10 and MB25 are shown Directly below these inserts, solid bars indicate the distribution of highly repetitive DNA sequences detected in the human c-*sis* region The next solid line represents a schematic restriction endonuclease map of the human c-*sis* containing DNA region In this map, v-*sis* homologous genetic sequences are indicated as heavy bars. Homology with either half of the v-*sis* regions is indicated. Open boxes represent the DNA inserts of pAO70, pAO79 and pAO121, which were used as molecular probes The DNA insert of pAO73 contains the 3′ v-*sis* homologous region and, therefore, is depicted as a heavy bar The next line represents the SSV proviral genome The open boxes are the long terminal repeats (LTRs) The 5′ and 3′ halves of the viral oncogene (v-*sis*) were subcloned in bacteriophage m13mp8 resulting in FAO1 and FAO2, respectively. The restriction endonuclease map of the feline c-*sis* proto-oncogene is shown under the SSV genome. Again, the genetic sequences homologous to v-*sis* are given as heavy bars Their hybridization with the DNA inserts of FAO1 and FAO2 is specified. Feline DNA fragments containing genetic sequences homologous to the human DNA inserts of pAO70 and pAO121 are indicated as open boxes. The solid bars below the feline restriction endonuclease map represent highly repetitive DNA sequences within the feline c-*sis* region. The two solid lines at the bottom of the figure indicate the size and the relative position of the feline cellular DNA inserts within cosmid clones MB65 and MB70. B, *Bam*HI, E, *Eco*RI, H, *Hin*dIII, K, *Kpn*I, Xb, *Xba*I, Xh, *Xho*I.

Comparative analysis of the human and feline c-sis locus

As an approach to identify c-*sis*-specific coding sequences outside the v-*sis* homologous regions, a comparative analysis of the cosmid clones obtained from the human and feline gene libraries was performed. First of all, highly repetitive genetic sequences were identified. This mapping was performed by Southern blot analysis using a number of restriction endonucleases and with ^{32}P-labeled total human or feline DNA as a molecular probe (data not shown). The position of the repetitive sequences within the human and feline c-*sis*-containing DNA region was determined and the results are summarized in Fig. 3. As can be seen, relatively long stretches of nonrepetitive DNA are present upstream of the v-*sis* homologous regions of both species. To see whether both species share homologous genetic sequences in this area, a number of restriction enzyme digests of the human c-*sis* cosmid clone were hybridized with ^{32}P-labeled feline c-*sis* containing cosmid clones (data not shown). The results indicated that there were homologous genetic sequences in addition to the v-*sis* related sequences. These sequences were found in an *Eco*RI/*Hin*dIII DNA fragment of 2.0 kbp (pAO121) and a *Kpn*I/*Kpn*I DNA fragment of 3.8 kbp (pAO70), subcloned and used as probes in further studies (see Fig. 3). Southern blot analysis of human and feline genomic DNA with these probes indicated that they both represented unique DNA sequences (data not shown). Hybridization

analysis of the feline c-*sis* locus with probe pAO121 revealed a 3.5 kbp *Hin*dIII/*Xho*I hybridizing DNA fragment (Fig. 4A, lane 1). Upon digestion with restriction endonucleases *Bam*HI, *Hin*dIII and *Xho*I a 2.0 kbp and an 1.5 kbp hybridizing fragment were seen (Fig. 4A, lane 2). Hybridization of the feline c-*sis* clones with pAO70 yielded another *Hin*dIII/*Xho*I restriction fragment of 3.5 kbp (Fig. 4B, lane 1). Upon digestion of the feline cosmid clones with restriction endonucleases *Bam*HI, *Hin*dIII and *Xho*I, a 3.5 kbp and an 1.5 kbp hybridizing DNA fragment were detected (Fig. 4B, lane 2) (for localization of the DNA fragments, see Fig. 3). In Fig. 3, the identified homologous regions in the human and feline c-*sis* clones have been depicted. These results indicate the presence of unique and conserved genetic sequences in close proximity and upstream of the v-*sis* homologous region in the c-*sis* proto-oncogenes of both species.

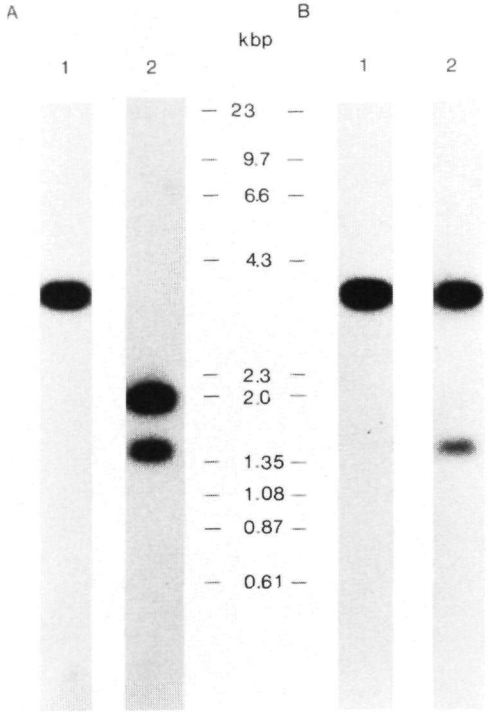


Fig. 4. Identification of unique genetic sequences in feline cosmid clone MB70 that are homologous to the human DNA inserts of pAO70 and pAO121. MB70 DNA (0.1 μg) was digested with restriction endonucleases *Hin*dIII and *Xho*I (A, lane 1; B, lane 1) or with *Bam*HI, *Hin*dIII and *Xho*I (A, lane 2; B, lane 2), electrophoresed through an 0.6% agarose gel, blotted onto nitrocellulose and hybridized to the ^{32}P-labeled 2.0 kbp *Eco*RI/*Hin*dIII insert of pAO121 (A) or 3.8 kbp *Kpn*I/*Kpn*I DNA insert of pAO70 (B). Molecular weight markers include λDNA digested with restriction endonuclease *Hin*dIII and DNA of φX174 digested with *Hae*III.

Northern blot analysis of human c-sis transcripts

The possibility that the additional homologous genetic sequences identified in the comparative analysis described above could represent a coding portion of the c-*sis* locus was tested in Northern blot analysis. Using a v-*sis*- (FAO1 and FAO2) and c-*sis*-specific (pAO73) probe, poly(A) selected RNA isolated from human lung tumor cells (lung carcinoma No. 4923) was shown to contain c-*sis* transcripts (Fig. 5, lanes 3, 4 and 5). No such transcripts could be detected in a similar analysis of RNA isolated from human cell line IARC-EW1 [43] (Fig. 5, lanes 1 and 2). With probe pAO121, the same 4.2 kb transcript as detected with the v-*sis*- and c-*sis*-specific probes appeared to hybridize (Fig. 5, lane 6). These results strongly favor the possibility that some genetic sequences in the

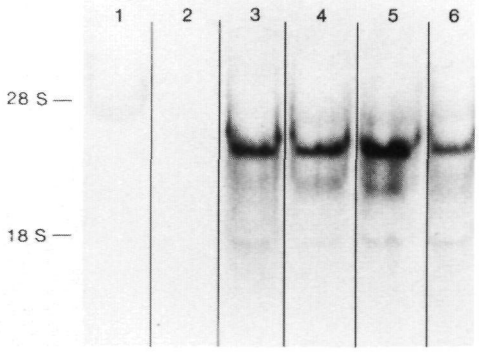


Fig. 5. Northern blot analysis of RNA transcripts in human lung tumor cells. Poly(A)-selected RNA samples were obtained from a human cell line (IARC-EW1) (lanes 1 and 2) and a human lung tumor (lung carcinoma No. 4923) (lanes 3, 4, 5 and 6) and screened (45-μg aliquots) in Northern blot analysis with the following *sis* probes: FAO1 (lanes 1 and 3), FAO2 (lanes 2 and 4), pAO73 (lane 5) and pA0121 (lane 6). Molecular weight markers included 28 S and 17 S ribosomal RNA.

2 0 kbp *Eco*RI/*Hin*dIII DNA fragment represent human c-*sis* coding sequences

Discussion

In the present study, the v-*sis* homologous DNA sequences of human and cat have been studied by Southern blot analysis followed by molecular cloning of both *sis* cellular homologs Southern blot analysis revealed that the v-*sis* homologous region in both species was rather extensive Furthermore, a *Hin*dIII-RFLP was frequently detected in the human locus Among human DNA samples from ten unrelated individuals, including five leukemic patients, six appeared to have an additional *Hin*dIII site in one allele The other four were homozygous for the absence of this site Interestingly, no individual was found to possess the *Hin*dIII site in both alleles The additional *Hin*dIII site could be localized in the middle of the v-*sis* homologous DNA region

With the v-*sis* viral oncogene as a molecular probe, the c-*sis* containing DNA regions were isolated from cosmid gene libraries of the two species A contiguous human DNA region of about 65 kbp was isolated and it appeared to contain all the human v-*sis* homologous sequences distributed discontinuously over a DNA region of approx 12 kbp The restriction endonuclease digestion data obtained in this study confirm data reported by Johnsson et al [41], Chiu et al [42] and Josephs et al [44,45] In these studies, the size and distribution of the six human v-*sis* homologous regions are described in detail At the 3′ end, the presence of an untranslated region was reported and part of it appeared to lack homology with v-*sis* [45] This particular region is also present in cosmid clone ALLW-1283-Cl21

The complete feline c-*sis* cellular homolog was represented in the inserts of cosmid clones MB65 and MB70 The feline v-*sis* homologous sequences were dispersed over approx 12 kbp of the 60 kbp contiguous feline DNA region isolated Upon comparative hybridization analysis of the human c-*sis* cosmid clone ALLW-1283-Cl21 and the two feline clones, the existence of common and highly conserved genetic sequences in addition to the v-*sis* homologous regions became apparent Northern blot analysis established that a transcript of the c-*sis* proto-oncogene of man contained one of these sequences These results indicate that the *sis* proto-oncogene contains as yet unidentified coding sequences that map upstream of the region that was captured in the viral oncogene of SSV Indications for the concept that the c-*sis* locus contains additional coding sequences and that SSV acquired only a portion of this gene could also be deduced from the size difference of human c-*sis* transcripts [26,27] and the v-*sis* oncogene [23,24] DNA sequence analysis of the v-*sis* homologous DNA region in the human locus revealed that the first exon lacked an initiation codon, again indicating the presence of additional exon sequences at the 5′ end of c-*sis*

A cDNA clone, designated pSM-1, was described by others [46] This clone was obtained from mRNA isolated from HUT-102 cells and represented a truncated form of the c-*sis* mRNA About 2 7 kbp of 3′ c-*sis* coding sequences appeared to be present in pSM-1 and an initiation codon was found about 64 bp upstream of the v-*sis* homologous region Transfection of pSM-1 into NIH-3T3 cells resulted in morphological transformation, while genomic DNA isolated from HUT-102 cells failed to do the same These authors speculated that the upstream c-*sis* genetic sequences might code for regulatory domains of PDGF, the putative gene product of the c-*sis* gene The newly discovered c-*sis* exon described in this report might be instrumental in verifying their speculation

Acknowledgements

We thank A J M Roebroek for helpful discussions and preparation of molecular probes This work was supported by the Netherlands Cancer Foundation, Contract No NUKC Bioch 80-1

References

1 Fishinger, P J (1982) in Molecular Biology of RNA Tumor Viruses (Stephenson J R, ed), pp 162–250 Academic Press New York
2 Bishop J M and Varmus, H (1982) in Molecular Biology of Tumor Viruses RNA Tumor Viruses (Weiss R, Teich, N Varmus, H E and Coffin J eds) pp 999–1108 Cold Spring Harbor Laboratory Cold Spring Harbor
3 Wong-Staal, F and Gallo, R C (1982) in Advances in Viral

Oncology (Klein, G, ed), pp 153–171, Raven Press, New York
4 Gonda, T J, Sheiness, D K and Bishop, J M (1982) Mol Cell Biol 2 617–624
5 Mueller, R, Slamon, D J, Tremblay, J M, Cline, M J and Verma, I M (1982) Nature 299, 640–644
6 Mueller, R, Verma, I M and Adamson, E D (1983) EMBO J 2, 679–684
7 Mueller, R, Tremblay, J M, Adamson, E D and Verma, I M (1983) Nature 304, 454–456
8 Cooper, G M (1982) Science 217, 801–806
9 Land, H, Parada, L F and Weinberg, R A (1983) Science 222, 771–778
10 Slamon, D J, deKernion, J B, Verma, I M and Cline, M J (1984) Science 224, 256–262
11 Yunis, J J (1983) Science 221, 227–236
12 Capon, D J Seeburg, P H, McGrath, J P, Hayflick, J S, Edman, U, Levinson, A D and Goeddel, D V (1983) Nature 304, 507–513
13 Maguire, R T, Robins, T S, Thorgeirsson, S S and Heilman, C A (1983) Proc Natl Acad Sci USA 80, 1947–1950
14 Fung, Y T, Lewis, W G, Crittenden, L B and Kung, H (1983) Cell 33 357–368
15 Collett, M S and Erikson, R L (1978) Proc Natl Acad Sci 75, 2021–2024
16 Van de Ven, W J M, Reynolds, F H, Jr and Stephenson, J R (1980) Virology 101, 185–197
17 Waterfield, M D, Scrace, G T, Whittle, N, Stroobant, P, Johnsson, A, Wasteson, A, Westermark, B, Heldin, C, Huang, J S and Deuel, T F (1983) Nature 304, 35–39
18 Doolittle, R F, Hunkapiller, M W, Hood, L E, Devare, S G, Robbins, K C, Aaronson, S A and Antoniades, H N (1983) Science 221, 275–277
19 Downward, J, Yarden, Y, Mayes, E, Scrace, G, Totty, N, Stockwell, P, Ullrich, A, Schlessinger, J and Waterfield, M D (1984) Nature 307, 521–527
20 Irgens, K, Wyers, M, Moraillon A, Parodi, A and Fortuny, V (1973) C R Acad Sci Paris 26, 1783–1786
21 Theilen, G H, Gould, D, Fowler M and Dungworth, D L (1971) J Natl Cancer Inst 47, 881–889
22 Besmer, P, Snyder, H W, Jr, Murphy, J E, Hardy, W D and Parodi, A (1983) J Virol 46, 606–613
23 Devare, S G, Reddy, E P, Law, J D, Robbins, K C and Aaronson, S A (1983) Proc Natl Acad Sci USA 80, 731–735
24 Devare, S G, Reddy, E P, Robbins, K C, Andersen, P R, Tronick, S R and Aaronson, S A (1982) Proc Natl Acad Sci USA 79, 3179–3182
25 Robbins, K C, Antoniades, H N, Devare, S G, Hunkapiller, M W and Aaronson, S A (1983) Nature 305, 605–608
26 Eva, A, Robbins, K C, Andersen, P R, Srinivasan, A, Tronick, S R, Reddy, E P, Ellmore, N W, Galen, A T, Lautenberger, J A, Papas, T S, Westin, E H, Wong-Staal, F, Gallo, R C and Aaronson, S.A (1982) Nature 295, 116–119
27 Westin, E H, Wong-Staal, F, Gelmann, E P, Dalla Favera, R, Papas, T S, Lautenberger, J A, Eva, A, Reddy, E P, Tronick, S R, Aaronson, S A and Gallo, R C (1982) Proc Natl Acad Sci USA 79, 2490–2494
28 Heldin, C, Westermark, B and Wasteson, A (1980) J Cell Physiol 105, 235–246
29 De Larco, J E and Todaro, G J (1978) Proc Natl Acad Sci USA 75, 4001–4005
30 Wong-Staal, F, Dalla Favera, R, Gelmann, E.P, Manzari, V, Szala, S, Josephs, S F and Gallo, R C (1981) Nature 294, 273–275
31 Gelmann, E P, Wong-Staal, F, Kramer, R A and Gallo, R C (1981) Proc Natl Acad Sci USA 78, 3373–3377
32 Ish-Horowicz, D and Burke, J F (1981) Nucleic Acids Res 9, 2989–2998
33 Messing, J and Vieira, J (1982) Gene 9, 269–276
34 Rigby, P W J, Dieckmann, M, Rhodes, C and Berg, P (1977) J Mol Biol 113, 237–251
35 Schalken, J A, Van den Ouweland, A M W, Bloemers, H P J and Van de Ven, W J M (1985) Biochim Biophys Acta 824, 104–112
36 Southern, E M (1975) J Mol Biol 98, 503–517
37 Verbeek, J S, Roebroek, A J M, Van den Ouweland, A M W, Bloemers, H P J and Van de Ven, W J M (1985) Mol Cell Biol 5, 422–426
38 Feramisco, J R, Helfman, D M, Smart, J E, Burridge, K and Thomas, G P (1982) J Biol Chem 257, 11024–11031
39 Steenbergh, P H, Hoeppener, J W M, Zandberg, J, Roos, B A, Jansz, H S and Lips, C J M (1984) J Clin Endocrinol Metab 58, 904–908
40 Dalla Favera, R, Gelmann, E P, Gallo, R C and Wong-Staal, F (1981) Nature 292, 31–35
41 Johnsson, A, Heldin, C, Wasteson, A, Westermark, B, Deuel, T F, Huang, J S, Seeburg, P H, Gray, A, Ullrich, A, Scrace, G, Stroobant, P and Waterfield, M D (1984) EMBO J 3, 921–928
42 Chiu, I, Reddy, E P, Givol, D, Robbins, K C, Tronick, S R and Aaronson, S A (1984) Cell 37, 123–129
43 Bechet, J M (1984) New Eng J Med 310, 393
44 Josephs, S F, Guo, C, Ratner, L. and Wong-Staal, F (1984) Science 223, 487–491
45 Josephs, S F, Ratner, L, Clarke, M F, Westin, E.H, Reitz, M S and Wong-Staal, F (1984) Science 225, 636–639
46 Clarke, M F, Westin, E, Schmidt, D, Josephs, S F, Ratner, L, Wong-Staal, F, Gallo, R C and Reitz, M S, Jr (1984) Nature 308, 464–467

CHAPTER 3

STRUCTURE AND NUCLEOTIDE SEQUENCE OF THE 5' REGION OF THE HUMAN AND FELINE *c-sis* PROTO-ONCOGENES.

Nucleic Acids Research 14 (1986), 765-778.

Structure and nucleotide sequence of the 5' region of the human and feline c-*sis* proto-oncogenes

Ans M W Van den Ouweland, Anton J M Roebroek, Jack A.Schalken[1], Chris A.A Claesen[2], Henri P J Bloemers and Wim J M Van de Ven

Molecular Oncology Section, Department of Biochemistry, University of Nijmegen, Geert Grooteplein Noord 21, 6525 EZ Nijmegen and Departments of [1]Urology and [2]Organic Chemistry, University of Nijmegen, Nijmegen, The Netherlands

Received 21 October 1985, Revised and Accepted 16 December 1985

ABSTRACT

Comparative analysis of cosmid clones containing the human and feline c-sis genetic regions revealed the similar structural organization of these areas in the two species. The areas shared seven different genetic regions in and around the c-sis locus and one of these was related to v-sis. Another region, 1.9 kbp in size and located about 8 kbp upstream of the v-sis homologous region in the human genome, also hybridized to the main c-sis transcriptional product of 3.5 kb. Comparison with a recently described c-sis cDNA clone (Collins et al., Nature 316, 748-750 (1985)) revealed that the 1.9 kbp DNA region contained a large 5' c-sis exon of at least 1050 bp. In this exon, the presumed initiation site of the predicted PDGF-2 containing precursor protein was located and appeared to be preceded by a large untranslated region. In the region immediately upstream of this exon, a TATA box and a consensus sequence for a potential Sp1 binding site were found at similar positions in both species. This region also exhibited promoter activity when tested in an assay in which coding sequences of bacterial chloramphenicol acetyltransferase (CAT; acetyl-CoA: chloramphenicol 3-O-acetyltransferase, EC 2.3.1.28) were placed under its control. The five other DNA regions were found upstream and downstream of the human c-sis transcription unit and also in an intron. Four of them contained repetitive sequences.

Hybridization analysis of human and feline c-sis containing cosmid clones with a mixed synthetic nucleotide probe, which corresponded to sequences encoding amino acid residues 2-7 of chain 1 of platelet-derived growth factor (PDGF-1), suggested that the c-sis cosmid clones did not include PDGF-1-specific genetic sequences.

INTRODUCTION

The human c-sis proto-oncogene contains coding sequences for a precursor protein, part of which is similar if not identical to chain 2 of platelet-derived growth factor (PDGF-2) (1,2). The initial link between PDGF-2 and sis was provided by partial amino acid sequence analysis of PDGF and nucleotide sequence analysis of the v-sis oncogene of simian sarcoma virus (3-5). Genetic sequences encoding PDGF-1 remain to be identified. Expression of c-sis is found in endothelial cells (6-8), human placenta (9), mouse embryo tissue (10), human lung tumors (11) and a number of tumor cell lines (12).

As the main transcriptional product, a mRNA species of about 3.5 kb (6,8,11,12), with coding potential for a protein of about 27 kd, was described (8,13). The precise function of the predicted c-sis encoded protein remains to be defined but it probably exhibits mitogen activity since it resembles PDGF with respect to amino acid sequence and also antigen and receptor binding characteristics (1,2,14-16). The responsiveness of fibroblasts, mesenchymal cells and glial cell lines to the mitogen PDGF and also the stimulation by PDGF of DNA synthesis during wound healing is well established (17-19).

It was recently reported that expression of the c-sis gene is modulated during endothelial cell differentiation in vitro (7). Furthermore, it was shown in transfection experiments, that expression in mouse embryo fibroblasts of human cDNA and genomic DNA constructs of c-sis resulted in a similar cell transformation (20,21) as observed upon transfection of the viral oncogene v-sis (22). Genetic sequences within or in close proximity of the locus may be involved in the control of such modulation of expression and they may also be relevant to the malignant potential of the locus. In an attempt to identify such regulatory sequences, we have compared the genetic organization of the c-sis gene and its flanking sequences in two different species, man and cat. In this report, we present the identification and characterization of seven genetic regions that human and cat share in and around the c-sis transcription unit.

MATERIALS AND METHODS

Cosmid and plasmid clones, E.coli strains and cell lines: Isolation of the human and feline c-sis cosmid clones was described previously (11). pAO68, pAO70, pAO73, pAO121, pAO151, pAO154 and pAO155 are subclones of the human c-sis locus (see also Fig. 2). pAO70, pAO73 and pAO121 were described elsewhere (11). pAO68 consists of a 3.8 kbp EcoRI/EcoRI DNA fragment subcloned in pSVBR91. pAO154 and pAO155 consist, respectively, of a 1.3 kbp BamHI/EcoRI and a 4.3 kbp BamHI/XhoI restriction fragment subcloned in pUC18. pAO151 consists of a 1.4 kbp EcoRI/HindIII restriction fragment subcloned in pAT153. pAO144 is a recombinant of pUC18 containing a 3.5 kbp HindIII/XhoI restriction fragment of the feline c-sis clone MB65. pSV2 was described by Gorman et al. (23) and pSuperCAT, a derivative of pSV2, was obtained from Dr. B. Dekker, University of Leiden, The Netherlands. pAO165 and pAO166 are recombinants of pSuperCAT containing an 0.4 kbp PstI/PstI human DNA fragment that is located immediately upstream of the 5′ c-sis exon

sequences. The insert in pAO166 has the same orientation relative to the chloramphenicol acetyltransferase (CAT) coding sequences as to the c-sis coding sequences in the human genomic DNA. The orientation of the insert in pAO165 is opposite to that in pAO166. For the DNA sequence analysis, the DNA fragments were inserted into the polylinker site of M13mp8-11 (24). E.coli strain HB101 was used as a host for the human and feline c-sis subclones. The recombinant M13 bacteriophages were propagated in E.coli strain JM101. Cell lines used in this study included HeLa (American Type Culture Collection (ATCC) CCL 2), Vero (ATCC CCL 81), HUT 102 (12) and Neuro 2A (ATCC CCL 131).

Preparation of DNA probes and hybridization: Preparation of DNA probes and their labeling by nick translation was carried out as described (25). The specific activity of the probes used in the hybridization studies was (2-5)×10^8 cpm/µg. Agarose gel electrophoresis, Southern blotting and hybridization analysis was performed as described (25). The mixed oligonucleotide probe was synthesized as described (26) and the probe was labeled with γ-[^{32}P]-ATP (3000 Ci/mmol) according to the procedure described by Maxam and Gilbert (27) to a specific activity of about 2×10^8 cpm/µg. Hybridization analysis with the mixed oligonucleotide probe was performed at 35 °C for 16 h in 6×SSC, 5×Denhardt's solution, 100 µg/ml denatured salmon testis DNA and 2×10^6 cpm/ml labeled mixed probe. Following hybridization, filters were washed in 5 mM EDTA pH 8.0 and 6×SSC during one hour periods at 37 °C, 40 °C, and 45 °C, successively, and dried. Autoradiography was performed by exposure to XAR-2 film (Kodak) with Dupont lightning plus intensifying screens.

DNA sequence analysis: Sequencing of DNA fragments was according to the dideoxy method of Sanger et al. (28). The gels readings were recorded, edited and compared using the Staden programs (29).

RNA isolation, Northern blotting and hybridization: Total cellular RNA was isolated using the lithium-urea procedure described by Auffray and Rougeon (30). Ten µg of mRNA, purified by oligo(dT)-cellulose affinity chromatography, was glyoxalated, fractionated on 1.0 % agarose gels (31) and transferred to Hybond-N (Amersham). Hybridization of the Northern blots was carried out as described by Church and Gilbert (32).

Assay for chloramphenicol acetyl transferase activity: The assay for chloramphenicol acetyltransferase (CAT) activity was performed as described by Gorman et al. (23).

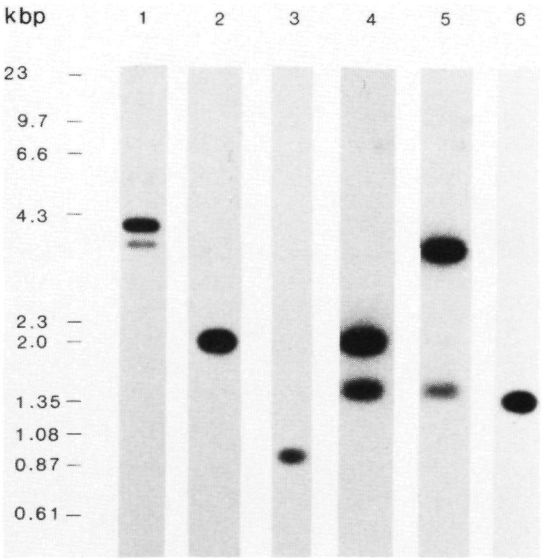


Fig. 1. Identification of common genetic sequences in the inserts of the human and feline c-sis cosmid clones. MB65 (lanes 1, 2 and 3) and MB70 (lanes 4, 5 and 6) were digested with restriction endonucleases EcoRI and KpnI (lane 1), BamHI and XbaI (lane 2), HindIII and BamHI (lane 3), HindIII, BamHI and XhoI (lanes 4 and 5) and EcoRI and XbaI (lane 6). As molecular probes, the inserts of the following human c-sis subclones were used: pAO155 (lane 1), pAO151 (lane 2), pAO154 (lane 3), pAO121 (lane 4), pAO70 (lane 5) and pAO68 (lane 6). Molecular weight markers include λ DNA digested with restriction endonuclease HindIII and DNA of ΦX174 digested with HaeIII.

RESULTS

Identification and characterization of seven genetic regions shared by the human and feline c-sis loci.

As an initial approach to identify potential regulatory regions within or in close proximity of the c-sis locus, we compared previously described human and feline c-sis containing cosmid clones (11) by Southern blot analysis. We reasoned that genetic regions which modulated expression of c-sis and which might be relevant to the malignant potential of the locus were likely to be conserved during evolution. A number of DNA fragments could indeed be identified when ^{32}P-labeled feline c-sis cosmid clones were hybridized with the human c-sis containing clones (data not shown). Following molecular cloning of 11 DNA fragments from the human c-sis containing cosmid clones and using them as probes, seven different genetic regions which were shared by man and cat could be identified (Fig. 1). One of these regions

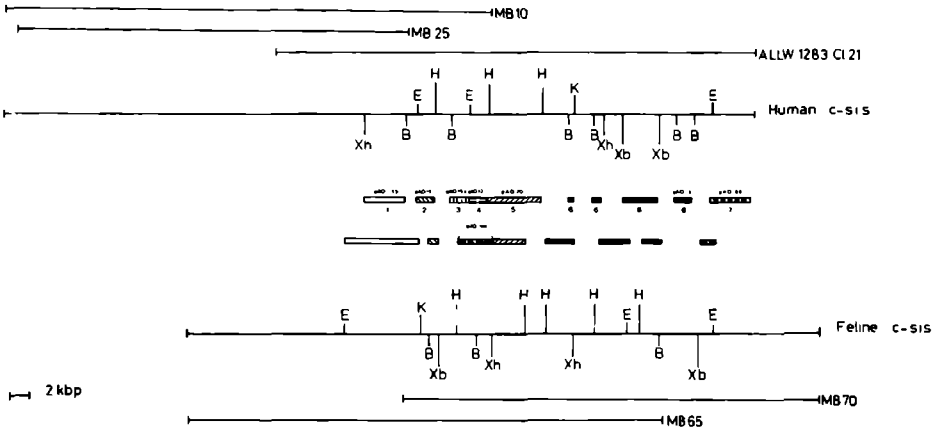


Fig. 2. Homologous regions in and around the human and feline c-sis locus. In the upper part of the figure, the size and relative positions of the DNA inserts of three human cosmid clones (ALLW-1283-C121, MB10 and MB25) are shown. Directly below these inserts, a schematic restriction endonuclease map of the human c-sis containing DNA region is presented. At the bottom of the figure, the size and relative positions of the cellular DNA inserts of two feline cosmid clones (MB65 and MB70) are given. Immediately above them, a schematic restriction endonuclease map of the feline c-sis containing DNA region is presented. Between the two restriction maps, the seven regions that man and cat share in this area are depicted as shaded boxes. Boxes of corresponding regions are shaded similarly and overlapping regions are shaded accordingly. The human regions are numbered (1-7). Human region 6, which represents the v-sis homologous region, is subdivided in four v-sis homologous fragments which all belong to region 6. The positions of the DNA inserts of recombinant plasmids are indicated by the names of the plasmids above the shaded boxes.
B, BamHI; E, EcoRI; H, HindIII; K, KpnI; Xb, XbaI; Xh, XhoI. (For a detailed restriction endonuclease map see Ref. 11).

(region 6) was homologous to v-sis (see also ref. 11) and a unique c-sis-specific probe (pA073) was prepared from the 3′ part of this region. In Fig. 2, the localization of the seven DNA regions within the human and feline c-sis containing genetic area is depicted.

To define these seven regions in and around the c-sis locus of the two species in more detail, we have further analyzed the molecular cloned human DNA fragments. It appeared from Southern blot analysis that region 1, 2, 3, 6 and 7 contained repetitive sequences (data not shown). The presence of Alu-repeats in region 6 (33) was confirmed. In region 1, 2, 3 and 7, no hybridization was observed when total human DNA was used as a probe. The repetitive sequences in these regions, therefore, do not belong to the

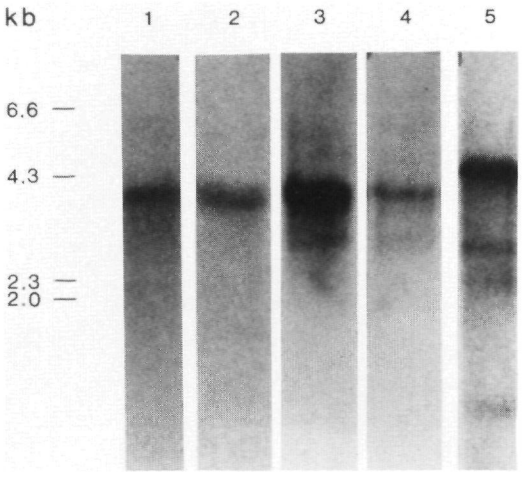


Fig. 3. Identification of c-sis transcripts in a number of cell lines. Poly(A)-selected RNA was isolated from HeLa (lanes 1 and 2), Vero (lane 3), HUT 102 (lane 4) and Neuro 2A (lane 5) cells and screened (10 µg aliquots) in Northern blot analysis with the inserts of pA0121 (lane 1), pA073 (lanes 2, 3, 4 and 5) as molecular probes. Molecular weight markers include λ DNA digested with the restriction endonuclease HindIII.

category of highly repetitive sequences (data not shown). Region 4 contained only unique genetic sequences. In Northern blot analysis of poly(A)-selected RNA isolated from HeLa cells, a 3.5 kb transcript was observed using pA073 (c-sis-specific) or pA0121 (containing region 4 as insert) as molecular probes (Fig. 3, lanes 1 and 2). Transcripts of similar size were also detected with these probes in a HTLV-I infected human T cell line (HUT102) (Fig. 3, lane 4) and the Vero cell line (Fig. 3, lane 3). The c-sis-related transcript in murine neuroblastoma cells was slightly larger (3.4) (Fig. 3, lane 5). Region 5 also contained only unique DNA sequences but it did not hybridize to the 3.5 kb c-sis transcript (data not shown).

DNA sequence analysis of human region 4 and its feline counterpart.

To resolve the question to what extent the genetic sequences in the unique human region 4 and the corresponding region in the cat genome were homologous, we have performed DNA sequence analysis. The complete nucleotide sequence of human region 4 and most of its feline counterpart were compared (Fig. 4). A remarkable homology between the human and feline DNA could be observed. Nucleotides printed in capitals represent sequences also found in a recently published cDNA clone of human c-sis expressed in endothelial

cells (8) and they can therefore be considered as exon sequences. As can be seen in Fig. 4, a continuous stretch of at least 1050 nucleotides constitute a single new exon. Homology between man and cat in this exon was 90 %. Together with the six already established c-sis exons (13) this new exon could form a mRNA transcript of about 3.5 kb, a size similar to that of the main c-sis transcript observed in a number of human cells (6,8,11,12) (see also Fig. 3). At its 3' end, a consensus splice junction was found. The 5' end of the exon remains to be defined. Within the human and feline exon sequences, four initiation codons were present. Three of them did not seem to be functional since they were followed almost immediately by a stop codon. The fourth ATG codon, which was found 63 nucleotides upstream of the 3' splice junction, is in-phase with the PDGF-2 open reading frame in the following exons (8,13). The 63 bp coding region of the new exon contained genetic sequences for a hydrophobic amino terminal sequence of the predicted precursor protein of PDGF-2.

Within the 5' portion of the human exon, a triplet of repeats (GCAGCTC) was followed by a doublet of a similar repeat (GCAGCCC). Interestingly, one of the repeats of the triplet was deleted in the c-sis cDNA clone published by Collins et al. (8). Five other differences, all point mutations, between human region 4 and this c-sis cDNA clone could be observed, namely two single nucleotide substitutions (an A instead of a G at position 683 and a T instead of an C at position 1383), two deletions (position 580 and 1151) and one insertion (a C at position 1262) (Fig. 4).

The exon sequences in human region 4 were preceded by a TATAAA sequence, the consensus promoter sequence for RNA transcription of eucaryotic genes (Fig. 4) (35). At the same position in the feline DNA, the same TATAAA sequence was found (Fig. 4). The observation that the 3.5 kb c-sis transcripts observed in a number of cells have the same length as the combined overlapping sequences in a number of recently published cDNA clones (8,13) is in support of the possibility that transcription of the c-sis locus starts in close proximity of this TATAAA containing region. In addition to the TATAAA consensus sequence, some other promoter-specific sequences were present upstream of the TATA box (see Discussion).

To test this potential promoter region for promoter activity, the 0.4 kbp PstI/PstI human DNA fragment, which contained the TATAAA box and the other promoter-specific sequences, was subcloned in pSuperCAT and the recombinant pA0166 was transfected into HeLa cells. As negative controls, pSuperCAT itself or pA0165 (pSuperCAT containing the 0.4 kbp PstI/PstI DNA fragment in

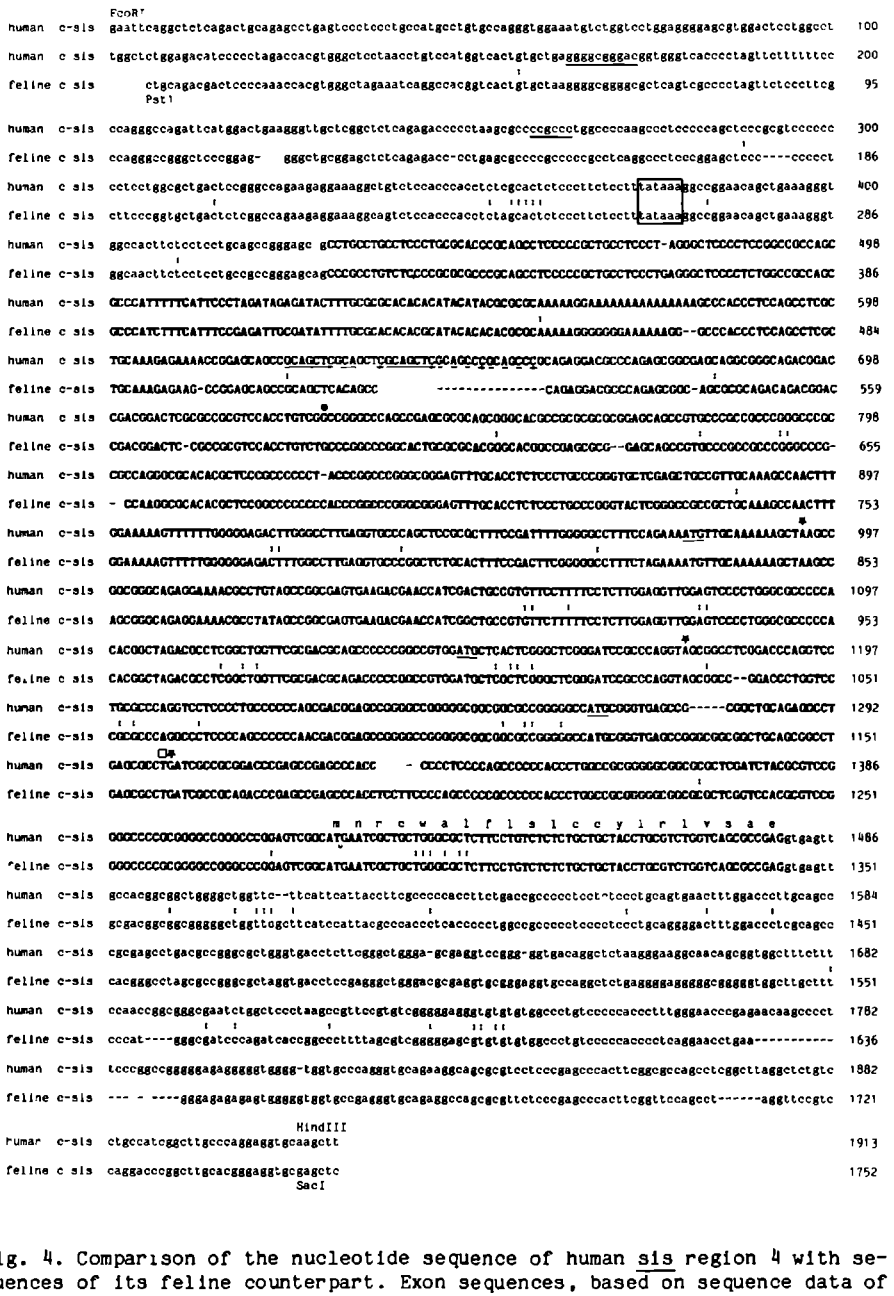


Fig. 4. Comparison of the nucleotide sequence of human sis region 4 with sequences of its feline counterpart. Exon sequences, based on sequence data of the sis cDNA clone isolated by Collins et al. (8), are printed in capitals. The first three ATG codons are underlined and asterisks (*) mark termination codons that are in-phase with a preceding ATG. The amino acid sequence of

the 63 bp open reading frame is given in the conventional one letter code. The putative TATA boxes of the human and feline locus are placed in a box. A closed circle (●) and an open square (□) mark the 5′ ends of, respectively, the human genomic c-sis subclone isolated by Gazit et al. (21) and the cDNA clone described by Ratner et al. (13). Two sets of small repetitive sequences are indicated with arrows (⟶,---⟶). The consensus sequence for the potential Sp1 binding sites and the CCGCCC sequence similar to the one found in the SV40 early promoter region are underlined.

an orientation opposite to that in pAO166) was used. As a positive control, pSV2 was included in the experiments (23). Measurements of CAT activity in HeLa extracts were made 48 hr after introduction of the DNA into the HeLa cells. Analysis of thin layer chromatograms revealed that there was a considerable amount of CAT enzymatic activity in extracts of HeLa cells transfected with pAO166. No such activity was detected in extracts from cells transfected with pSuperCAT or pAO165. These results indicated that the 0.4 kbp PstI/PstI human DNA fragment had the capacity to function as a promoter.

Analysis of the human and feline c-sis containing cosmid clones for the presence of PDGF-1 related sequences.

The observation that a number of DNA regions in and around the human and feline c-sis coding region, which encodes a protein similar if not identical to PDGF-2, appeared to be conserved, raised the possibility of the presence of genetic sequences related to those encoding PDGF-1. To test this possibility, we prepared a mixed synthetic oligonucleotide probe based upon the amino-terminal amino acid sequences (residues 2-7) (3,4,36) of PDGF-1. The composition of the synthetic probe is depicted in Fig. 5A. Hybridization analysis was positive in the case of a 5.0 kbp HindIII/HindIII DNA fragment of cosmid clone ALLW-1283-C121 (Fig. 5B, lane 3). None of the other human (Fig. 5B, lanes 1 and 2) or feline (data not shown) cosmid clones hybridized to the mixed synthetic probe. Upon cloning and further analysis of the 5.0 kbp HindIII/HindIII DNA fragment, the matching sequences were localized in a 1.0 kbp Sau3A/HindIII DNA fragment (Fig. 5C). Nucleotide sequence analysis of this fragment revealed a match of 15 out of 17 nucleotides of one of the probes in the mixture (Fig. 5D). However, this sequence appeared to be no part of a region encoding the amino-terminal portion of PDGF-1.

DISCUSSION

In this report, we describe the structure and the nucleotide sequence of the 5′ region of the human and feline c-sis proto-oncogenes and provide the

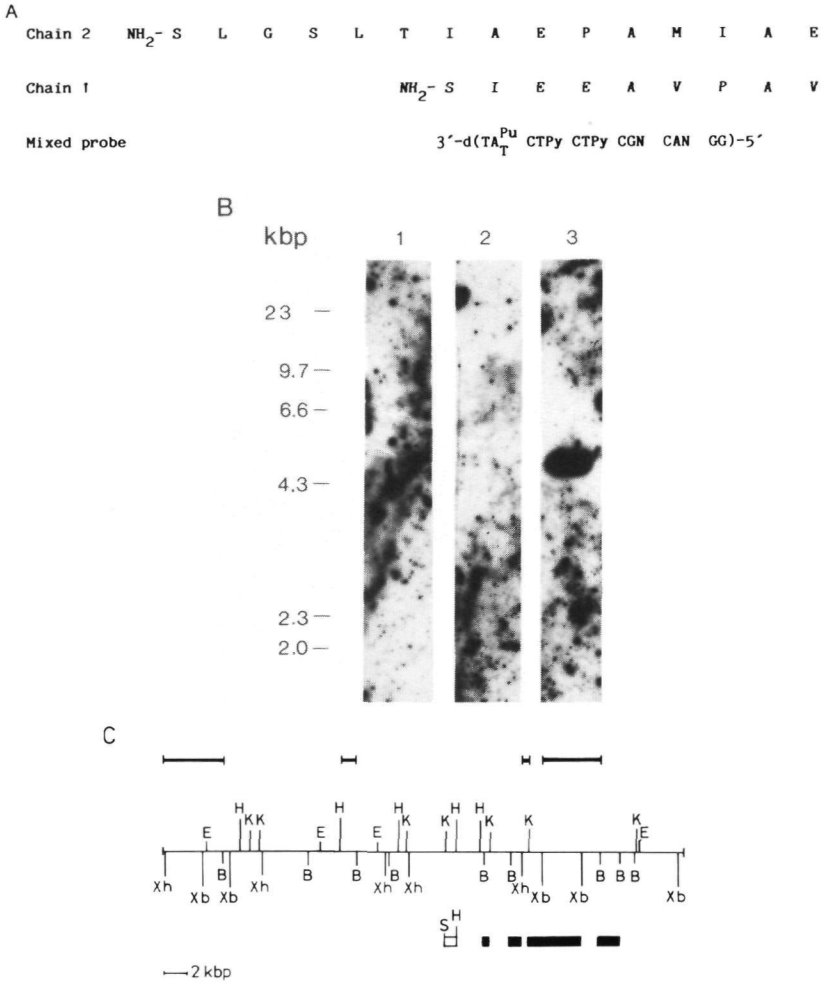


Fig. 5. Identification of genetic sequences homologous to the synthetic mixed probe. (A) Composition of the mixed probe. The N-terminal 15 amino acids of PDGF-2 and 9 amino acids of PDGF-1 are aligned for maximum homology (3,36). The composition of the mixed probe is indicated under the amino acid sequence data. (B) Southern blot analysis of human c-sis cosmid clones with the mixed probe. Clones MB10 (lane 1), MB25 (lane 2) and ALLW-1283-C121

(lane 3) were digested with the restriction endonuclease HindIII. Hybridization of the Southern blots was as described under Materials and Methods. (C) Restriction endonuclease map of the human c-sis cosmid clone ALLW-1283-C121. The upper heavy bars represent highly repetitive DNA sequences. The next line represents a schematic restriction endonuclease map of the human c-sis locus. At the bottom of the figure, black boxes indicate the relative positions of v-sis homologous sequences and the open box the DNA region that hybridized with the mixed synthetic probe.
B, BamHI; E, EcoRI; H, HindIII; K, KpnI; S, Sau3A; Xb, XbaI; Xh, XhoI. (D) Nucleotide sequence of the first 107 bp of the Sau3A/HindIII DNA fragment illustrating the homology with the mixed synthetic probe.

first characteristics of their promoter regions. Comparative analysis of the human and feline c-sis containing DNA areas resulted in the identification of seven DNA regions which the two species have in common. These regions are highly homologous in the two species and their genetic organizations are colinear. One of the regions, region 6, constitutes the human and feline v-sis cellular homolog which we have described before (11). In region 4, the 5′ region of the transcription unit of the c-sis proto-oncogene is located. We conclude this from the following observations. Combination of the overlapping inserts of the human sis cDNA clones (8,13) constitutes a stretch of DNA of about 3.5 kbp which resembles the size of the c-sis mRNA species detectable in a number of cell lines. This means that the large exon described in this paper contains almost all the 5′ genetic sequences of the messenger although the precise position of the cap site remains to be identified. Further support for our conclusion is based upon the fact that a TATA box is present in the DNA region immediately upstream of the large exon. Furthermore, this region exhibited promoter activity in an assay for transcriptional activity. The TATAAA consensus promoter sequence for RNA transcription of eukaryotic genes is present at the same positions in the c-sis loci of the two species. Since the distance in many eukaryotic genes between the TATA box and the cap site of the mRNA is about 30 nucleotides, it can be concluded that from the 5′ sis cDNA clone described by Collins et al. (8) only a few nucleotides of the 5′ region of the mRNA are excluded. In both species, additional promoter-specific consensus sequences are found upstream of the TATA box. A (G+C)-rich region is located about 80 nucleotides upstream of the putative RNA transcriptional starting region and the characteristic CCGCCC sequence (position 259-264 in Fig. 4) is present within this region. In the SV40 early promoter region, six copies of this sequence were found in a similar (G+C)-rich region (37) and in the myc proto-oncogene the hexanucleotide is located in close proximity to the TATA box (38).

At about 200 nucleotides upstream of the TATA box a GC box consensus sequence (GGGGCGGGAC in the human and GGGGCGGGGC in the feline DNA) for the binding site of the promoter-specific factor Sp1 (39) is present. A number of GC box-containing promoters have been described (39) and they include cellular as well as viral genes. However, it remains to be established whether the GC box in the sis locus is involved in Sp1 binding.

A striking characteristic of the 5' c-sis exon is its long noncoding region (at least 987 bp). The messenger of the insulin-like growth factor II precursor has a similar long 5' untranslated region (40). It is possible that the long 5' untranslated region of the c-sis locus plays a regulatory role but this needs further studies. The presence of four ATG initiation codons was also remarkable. Most likely, translation starts at the fourth initiation site (position 1417). This ATG codon meets the criteria postulated by Kozak (41,42) and it also is the beginning of an open reading frame of 723 bp (8,13) that encodes the predicted PDGF-2 precursor protein of about 27 kd.

The other five genetic regions which man and cat have in common in and around the c-sis transcription unit are not yet fully characterized. It is clear from our data that at least one of these regions lies in an intron and that its sequences are more conserved than other intron regions. This intron region is not present in the main c-sis transcript (data not shown). Furthermore, it is of interest to note that some of these regions contain repetitive DNA sequences. The presence of such repetitive sequences in the proximity of cellular coding sequences was reported before (43) and could have functional implications.

A still unresolved issue pertains to where the genetic sequences that encode the PDGF-1 chain are located. We have tested the possibility that they are present in the DNA areas described in this study. Southern blot analysis with a PDGF-1-specific synthetic mixed probe did not reveal any potential region. However, it should be noted that in the case that a splice site would be located in the sequences recognized by the mixed probe, a match could have escaped detection. Our results suggest that the PDGF-1 encoding sequences are located somewhere else, even on another chromosome.

ACKNOWLEDGEMENTS

We thank Carla Onnekink, Jan van Groningen and Anja ter Avest for their excellent technical assistance and Lambert Dorssers for helpful discussions. This work was supported by the Netherlands Cancer Foundation, Contract No. NUKC Bioch. 85-9.

REFERENCES

1. Chiu, I.M., Reddy, E.P., Givol, D., Robbins, K.C., Tronick, S.R. and Aaronson, S.A. (1984) Cell 37, 123-129.
2. Johnsson, A., Heldin, C.H., Wasteson, A., Westermark, B., Deuel, T.F., Huang, J.S., Seeburg, P.H., Gray, A., Ullrich, A., Scrace, G., Stroobant, P. and Waterfield, M.D. (1984) EMBO J. 5, 921-928.
3. Doolitle, R.F., Hunkapillar, M.W., Hood, L.E., Devare, S.G., Robbins, K.C., Aaronson, S.A. and Antoniades, H.N. (1983) Science 221, 275-277.
4. Waterfield, M.D., Scrace, G.T., Whittle, N., Stroobant, P., Johnsson, A., Wasteson, A., Westermark, B., Heldin, C.H., Huang, J.S. and Deuel, T.F. (1983) Nature 304, 35-39.
5. Robbins, K.C., Antoniades, H.N., Devare, S.G., Hunkapiller, M.W. and Aaronson, S.A. (1983) Nature 305, 605-608.
6. Barrett, T.B., Gajdusek, C.M., Schwartz, S.M., McDougall, J.K and Benditt, E.P. (1984) Proc. Natl. Acad. Sci. USA 81, 6772-6774.
7. Jaye, M., McConathy, E., Drohan, W., Tong, B., Deuel, T. and Maciag, T. (1985) Science 228, 882-885.
8. Collins, T., Ginsburg, D., Boss, J.M., Orkin, S.H. and Pober, J.S. (1985) Nature 316, 748-750.
9. Goustin, A.S., Betsholtz, C., Pfeifer-Ohlsson, S., Persson, H., Rydnert, J., Bywater, M., Holmgren, G., Heldin, C.H., Westermark, B., Ohlsson, R. (1985) Cell 41, 301-312.
10. Slamon, D.J. and Cline, M.J. (1984) Proc. Natl. Acad. Sci. USA 81, 7141-7145.
11. Van den Ouweland, A.M.W., Breuer, M.L., Steenbergh, P.H., Schalken, J.A., Bloemers, H.P.J. and Van de Ven, W.J.M. (1985) Biochim. Biophys. Acta 825, 140-147.
12. Eva, A., Robbins, K.C., Andersen, P.R., Srinivasan, A., Tronick, S.R., Reddy, E.P., Ellmore, N.W., Galen, A.T., Lautenberger, J.A., Papas, T.S., Westin, E.H., Wong-Staal, F., Gallo, R.C. and Aaronson, S.A. (1982) Nature 295, 116-119.
13. Ratner, L., Josephs, S.F., Jarrett, R., Reitz, M.S. and Wong-Staal, F. (1985) Nucl. Acids Res. 13, 5007-5018.
14. Betsholtz, C., Westermark, B., Ek, B. and Heldin, C.H. (1984) Cell 39, 447-457.
15. Nister, M., Heldin, C.H., Wasteson, A. and Westermark, B. (1984) Proc. Natl. Acad. Sci. USA 81, 926-930.
16. Stiles, C.D. (1983) Cell 33, 653-655.
17. Ross, R., Glomset, J., Kariya, B. and Harker, L. (1974) Proc. Natl. Acad. Sci. USA 71, 1207-1210.
18. Scher, C.D., Shepard, R.C., Antoniades, H.N. and Stiles, C.D. (1979) Biochim. Biophys. Acta 560, 217-241.
19. Heldin, C.H., Westermark, B. and Wasteson, A. (1980) J. Cell. Physiol. 105, 235-246.
20. Clarke, M.F., Westin, E., Schmidt, D., Josephs, S.F., Ratner, L., Wong-Staal, F., Gallo, R.C. and Reitz, M.S., Jr. (1984) Nature 308, 464-467.
21. Gazit, A., Igarashi, H., Chiu, I.M., Srinivasan, A., Yaniv, A., Tronick, S.R., Robbins, K.C. and Aaronson, S.A. (1984) Cell 39, 89-97.
22. Hannik, M. and Donoghue, D.J. (1984) Science 226, 1197-1199.
23. Gorman, C.M., Merlino, G.T., Willingham, M.C., Pastan, I. and Howard, B.H. (1982) Proc. Natl. Acad. Sci. USA 79, 6777-6781.
24. Messing, J. and Vieira, J. (1982) Gene 19, 269-276.
25. Schalken, J.A., Van den Ouweland, A.M.W., Bloemers, H.P.J. and Van de Ven, W.J.M. (1985) Biochim. Biophys. Acta 824, 104-112.
26. Claesen, C., Tesser, G.I., Dreef, C.E., Marugg, J.E., Van der Marel, G.A. and Van Boom, J.H. (1984) Tetrahedron Letters 25, 1307-1310.

27. Maxam, A.M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560-564.
28. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
29. Staden, R. (1982) Nucl. Acids Res. 10, 2951-2961.
30. Auffray, C. and Rougeon, F. (1980) Eur. J. Biochem. 107, 303-314.
31. McMaster, G.K. and Carmichael, G.G. (1977) Proc. Natl. Acad. Sci. USA 74, 4835-4838.
32. Church, G.M. and Gilbert, W. (1984) Proc. Natl. Acad. Sci. USA 81, 1991-1995.
33. Favera, R.D., Gelmann, E.P., Gallo, R.C. and Wong-Staal, F. (1981) Nature 292, 31-35.
34. Van Zoelen, E.J.J., Van de Ven, W.J.M., Franssen, H.J., Oosterwaard, T.M.J., Van der Saag, P.T., Heldin, C.H. and De Laat, S.W. (1985) Mol. Cell. Biol. 5, 2289-2297.
35. Breathnach, R. and Chambon, P. (1981) Ann. Rev. Biochem. 50, 349-383.
36. Antoniades, H.N. and Hunkapiller, M.W. (1983) Science 220, 963-965.
37. Fromm, M. and Berg, P. (1982) J. Mol. Appl. Genet. 1, 457-481.
38. Ishii, S., Xu, Y.H., Stratton, R.H., Roe, B.A., Merlino, G.T. and Pastan, I. (1985) Proc. Natl. Acad. Sci. USA 82, 4920-4924.
39. Dynan, W.S. and Tjian, R. (1985) Nature 316, 774-778.
40. Dull, T.J., Gray, A., Hayflick, J.S. and Ullrich, A. (1984) Nature 310, 777-781.
41. Kozak, M. (1984) Nucl. Acids Res. 12, 857-872.
42. Kozak, M. (1984) Nature 308, 241-146.
43. Efstratiadis, A., Posakony, J.W., Maniatis, T., Lawn, R.M., O'Connell, C., Spritz, R.A., DeRiel, J.K., Forget, B.G., Weissman, S.M. Slightom, J.L., Blechl, A.E., Smithies, O., Baralle, F.E., Shoulders, C.C. and Proudfoot, N.J. (1980) Cell 21, 653-668.

CHAPTER 4

GENETIC ORGANIZATION OF THE c-*sis* TRANSCRIPTION UNIT

Nucleic Acids Research 15 (1987), 959-970.

Genetic organization of the c-*sis* transcription unit

Ans M W Van den Ouweland, Jan J M Van Groningen, Jack A Schalken[1], Han W Van Neck, Henri P J Bloemers and Wim J M Van de Ven

Molecular Oncology Section, Department of Biochemistry, University of Nijmegen, Geert Grooteplein Noord 21, 6525 EZ Nijmegen, The Netherlands and [1]Department of Urology, University of Nijmegen, The Netherlands

Received November 11 1986 Accepted January 8, 1987

ABSTRACT

The structure and genetic organization of the transcription unit of the c-sis proto-oncogene was determined. Comparative nucleotide sequence analysis of the exon sequences of feline and human c-sis revealed a very high degree of homology. The cap site as well as the poly(A)-addition site of the sis transcript of each species was identified and found in similar positions. An insert of 4 amino acids was found in the deduced translational product of feline c-sis and it was located at the amino-terminus of the region that constituted the platelet-derived growth factor domain. An insert of 149 bp present at the 5' end of exon 7 of human c-sis but absent in the simian sarcoma virus v-sis oncogene was also present in the feline c-sis proto-oncogene.

INTRODUCTION

Upon the demonstration that the function and the amino acid sequence of one chain of the human platelet-derived growth factor (PDGF) was nearly identical to the oncogene product of the simian sarcoma virus (SSV) (1,2,3), studies on the cellular c-sis counterpart got great impetus. Sequences of the c-sis proto-oncogene appeared also to be acquired by another acutely transforming retrovirus, namely the Parodi-Irgens strain of feline sarcoma virus (PI-FeSV) (4). The c-sis-derived sequences in PI-FeSV and SSV are expressed in different contexts of the retroviral genomes. PI-FeSV encodes a 76 kda gag-sis polyprotein (5) and SSV a 28 kda protein that contains a short NH_2-terminal leader sequence derived from the viral env gene (6,7). In both cases, constitutive expression of these PDGF-like mitogens seems capable of inducing neoplastic transformation (4,8).

In previous studies, the structure and genetic organization of the human (9,10,11) and feline (12,13) c-sis proto-oncogenes were described. Using nucleotide sequence analysis of genomic and cDNA clones, the distribution of the 7 human c-sis exons was obtained (14,15,16). In the region upstream of the first human sis exon, a "TATA" box and consensus sequences for a potential Sp1 binding site were found. At similar positions in the feline c-sis

proto-oncogene, the same sequences were present. The upstream region exhibited promoter-like activity when tested in an assay in which coding sequences for bacterial chloramphenicol-acetyl-transferase were placed under its control (13). Therefore, it possibly represents the promoter region of the c-sis locus and controls the synthesis of the c-sis mRNA (13). Further evidence for this possibility could be obtained by precisely defining the sis transcript. In the present report, we provide such evidence. We present data from nuclease S1 protection experiments and comparative nucleotide sequence analysis of sis cDNA clones and define the 5' and 3' boundaries of the feline and human sis transcripts.

MATERIALS AND METHODS

Biological materials: Cells were grown in Dulbecco's modification of Eagle's medium (DMEM) supplemented with 10% calf serum (GIBCO). Cell lines used in this study included HeLa (American Type Culture Collection (ATCC) CCL 2), ML-1 (17) and PC3 (ATCC CRL 1435). Feline placentas were obtained from the Central Animal Laboratory of the University of Nijmegen, Nijmegen.

Recombinant DNA clones: Isolation of human and feline c-sis cosmid clones and pA073, a human c-sis specific subclone, was described previously (12). pHN1 is a recombinant clone of λgt11 and a 0.6 kbp sis-specific cDNA isolated from an oligo(dT)-primed human cDNA library. A44, a subclone of pA0121 (12), consists of a 243 bp AluI/AluI DNA restriction fragment subcloned in M13mp11. A53 is a recombinant of M13mp11 containing a 235 bp AluI/AluI DNA restriction fragment isolated from pA0144 (13). A 1.3 kbp BamHI/HindIII DNA restriction fragment, which represents a 3'-specific portion of c-sis, was isolated from MB70 (12) and subcloned in pUC18 resulting in pPHS1. pAct is an actin-specific clone described by Dodemont et al. (18). E.coli strain HB101 was used for the propagation of plasmids and the recombinant M13 bacteriophages were propagated in E.coli strain JM109.

DNA sequence analysis: DNA fragments were ligated into the polylinker region of M13 mp8-11 (19). Sequencing of the DNA fragments was according to the dideoxy method of Sanger et al. (20). Gel readings were recorded, edited and compared using the Staden programs (21).

RNA isolation and hybridization: Total cellular RNA was isolated using the lithium-urea procedure described by Auffray and Rougeon (22). Ten µg of oligo(dT)-cellulose purified mRNA was glyoxalated, fractionated on 1% agarose gels (23) and transferred to Hybond-N (Amersham). Isolation of DNA probes, their nick-translation and hybridization of Southern blots were car-

ried out as described before (24). Hybridization of Northern blots was performed according to the method described by Church and Gilbert (25).

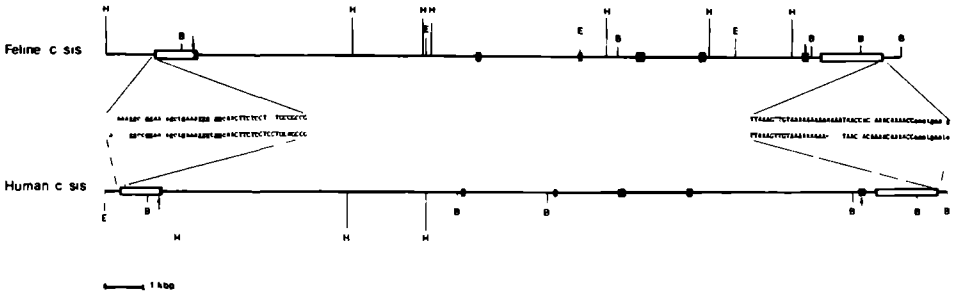
Construction and screening of cDNA libraries: Oligo(dT)-primed cDNA libraries of ML-1 or feline placenta mRNA were constructed in λgt11 as described by Roebroek et al. (26). About 1.2×10^6 plaques obtained upon infection of E.coli Y1090 (27), were screened as described by Hanahan and Meselson (28).

Nuclease S1 analysis: Nuclease S1 mapping was carried out as described by Van Leen et al. (29). Oligo(dT)-selected mRNA (5 µg) from feline placenta or human PC3 cells were hybridized with the inserts of A53 and A44, respectively. The inserts of A53 and A44 were labeled using the method of Sanger (20) and purified by electrophoresis in a 6% polyacrylamide/7 M ureum gel. Hybridizations were performed at 55°C. Nuclease S1 (Boehringer) concentrations used in the experiments were 25 units/µl and incubations were performed at 37°C.

RESULTS AND DISCUSSION

Genetic organization of feline c-sis.

In an approach to define the c-sis transcription unit, we decided not to limit our studies to the sis transcript of only the human species but to include also the feline sis transcript. We reasoned that by such an extension species-specific characteristics could be eliminated and a more general description of the c-sis transcription unit could be obtained. Detailed genomic information about the c-sis locus, however, was only available for human c-sis. To obtain the necessary genomic data about the feline c-sis locus, the genomic organization of the feline sis transcription unit was determined by comparative analysis of previously described cosmid clones (12,13). Overlapping inserts of these clones contained the complete feline or human c-sis proto-oncogene (12,13). Based upon the known exon distribution in human c-sis (15,16), the location of the feline c-sis exons was determined accurately by Southern blot analysis of the inserts of the feline cosmid clones using human DNA fragments containg c-sis exons as molecular probes (data not shown). In control experiments, DNA fragments identified as putative sis exons were tested for their ability to detect c-sis transcripts in Northern blot analyses (data not shown). Upon identification of DNA fragments that contained the feline c-sis exons, the nucleotide sequence of them was determined. The topological distribution of the feline and human c-sis exons is given in Fig. 1 and reveals great similarity. In Fig. 2,

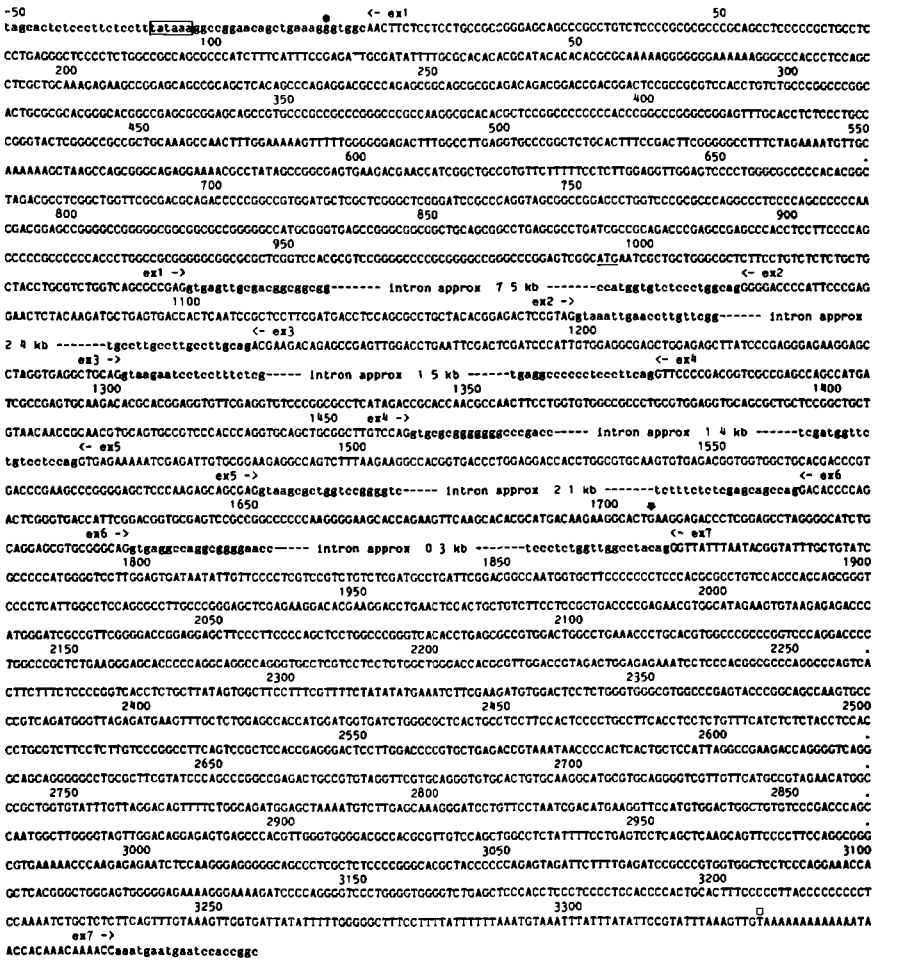


1. Topology of the feline and human c-sis exons. The bars in the upper and lower part of the figure represent schematic restriction endonuclease maps of the feline and human c-sis proto-oncogene, respectively. Open boxes represent the 5′ and 3′ noncoding exon sequences, while coding sequences are depicted as solid boxes. The positions of the used ATG (↓) and TGA (↓) codons are also shown. Between the restriction endonuclease maps, the nucleotide sequence of the 5′ and 3′ ends of the feline and human sis transcription units as well as their flanking sequences are given. Exon sequences are depicted as capitals.

nucleotide sequence data of the feline sis exons and a small portion of their flanking sequences are shown. It should be noted that results obtained from nucleotide sequence analysis of feline sis-specific cDNA clones that are described later in this report are already included in the figure. The position of the "TATA" consensus promoter sequence is also indicated in the figure.

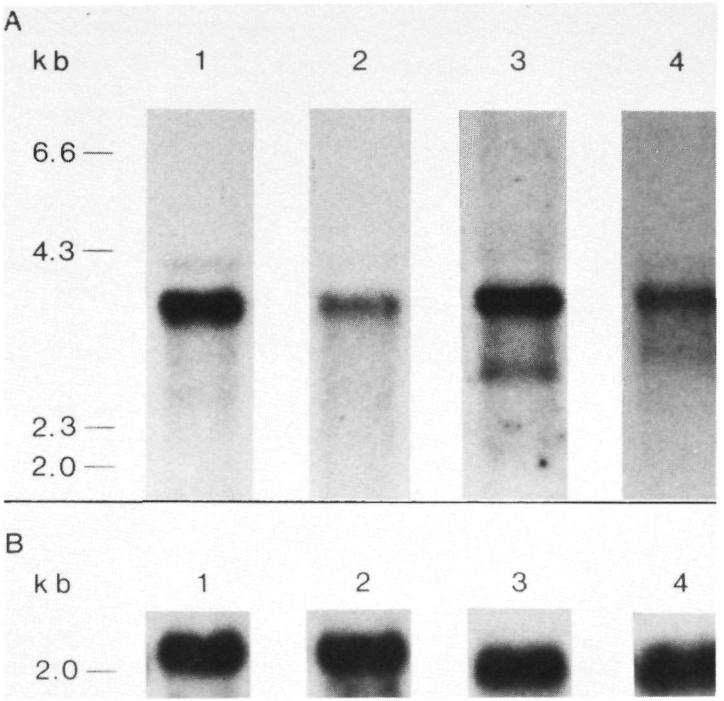
Expression of the feline c-sis proto-oncogene.

To precisely define the exons in the sis transcription unit, analysis of cDNA clones prepared with sis mRNA was required. To select appropriate sources for sis mRNA, we analyzed poly(A)-selected RNA from a human prostate carcinoma (PC3) and a myelomonoblastoma (ML-1) cell line and from specimens of feline placenta (Fig. 3A). Selection of feline placenta was based upon studies by Goustin et al. (30) who demonstrated the presence of relatively high levels of c-sis mRNA in placenta tissue. As a c-sis-specific molecular probe, pPHS1 was used. In an attempt to compare similar amounts of poly(A)-selected RNA, RNA concentrations were estimated based upon $O.D_{260}$ readings and ethidium bromide staining. As an additional control experiment, levels of actin transcripts were estimated by Northern blot analysis (Fig. 3B). Northern blot analysis of mRNA isolated from feline term placentas (Fig. 3A, lane 1) or from placentas containing embryos that were about 3 weeks old (Fig. 3A, lane 2) revealed the presence of a 3.5 kb sis-specific transcript. However, a difference in the level of expression of sis seemed to exist in



2. Nucleotide sequence of the feline c-sis proto-oncogene. Lower case letters indicate nucleotides not belonging to the exon sequences of the feline c-sis proto-oncogene. The exon sequences are presented by capitals. The "TATA" box sequences are placed in a box. A closed circle (●) and an open square (□) mark respectively the cap site and poly(A)-addition site determined by Rao et al. (16) in the human counterpart. The ATG codon preceeding the large open reading frame is underlined and an asterisk (*) marks the used TGA stopcodon.

the placenta specimens of the two stages of pregnancy. The amount of c-sis mRNA transcripts in term placenta was clearly higher than in the placenta with embryos of 3 weeks old. As can be seen in Fig. 3A, the sis mRNA level in PC3 cells (lane 3) was somewhat higher than in ML-1 cells (lane 4). Com-



3. Northern blot analysis of RNA transcripts of feline placentas and human cell lines. (A) Poly(A)-selected RNA was isolated from feline term placentas (lane 1), from placentas containing 3 weeks old embryos (lane 2), human PC3 (lane 3) or human ML-1 (lane 4) cells and screened with the inserts of pPHS1 (lanes 1 and 2) and pA073 (lanes 3 and 4) as molecular probes. (B) Screening of the same Northern blot as in (A) with the insert of pAct as molecular probe. Molecular weight markers include λ DNA digested with the restriction endonuclease HindIII.

parison of the size of the human and feline c-sis mRNAs revealed no major differences. A weak hybridization signal at about 2.6 kb was detected in RNA preparations from PC3 and ML-1 cells. The nature and origin of this transcript is not known at present but it was not seen in the feline placenta specimens. Slamon and Cline (31) have studied sis expression during embryonic development of the mouse. Higher levels of sis mRNA were found at early stages of development than at later stages. However, it should be noted that tissue specimens taken during the early stages of embryonic development included embryonic as well as extra-embryonic tissue. This in contrast to specimens from later stages in development in which only embryos were studied. The observed differences in the levels of sis transcription is of

interest, especially in light of the report that mRNA levels of the PDGF-receptor are higher in term placenta than in earlier stages of pregnancy (32). However, for a proper evaluation of the observations, identification of the cell types that express the PDGF receptor and the sis product is required.

Isolation of sis cDNA clones.

Using poly(A)-selected mRNA from feline term placenta or human ML-1 cells, a feline and human cDNA library was constructed in λgt11. Initial screening of the feline cDNA library (about 1×10^6 plaques) was performed with a combination of two different probes. These included a 3'-sis-specific probe (insert of pPHS1) and a 5'-sis-specific probe (a 1.5 kbp BamHI/XhoI insert of pAO144, which contained the 3'end of the first sis exon). Screening resulted in the identification of 16 positive cDNA clones. Of these, 15 hybridized to the 3'-sis-specific probe. Since the internal EcoRI restriction endonuclease recognition sites within the cDNAs were not methylated during the construction of the cDNA libraries, the inserts of all 15 clones appeared to be located downstream of the internal EcoRI site in exon 3 of the feline c-sis proto-oncogene. Only 1 positive clone (FAO184) was isolated whose insert was located upstream of the internal EcoRI site. The nucleotide sequence of the cDNA insert of FAO184 and of 6 cDNAs that were located downstream of the internal EcoRI recognition site was determined. The insert of FAO184 starts at position 835 and stops at the EcoRI recognition site. The inserts of the 6 other cDNAs varied in lenght between 0.9 and 2.2 kbp. One cDNA contained a poly(A) tail and, based upon this observation, the poly(A)-addition site could be mapped at position 3355 (Fig. 2). The 3' ends of the other five cDNAs stopped at various positions (3301; 3301; 3327; 3332; 3343) (data not shown).

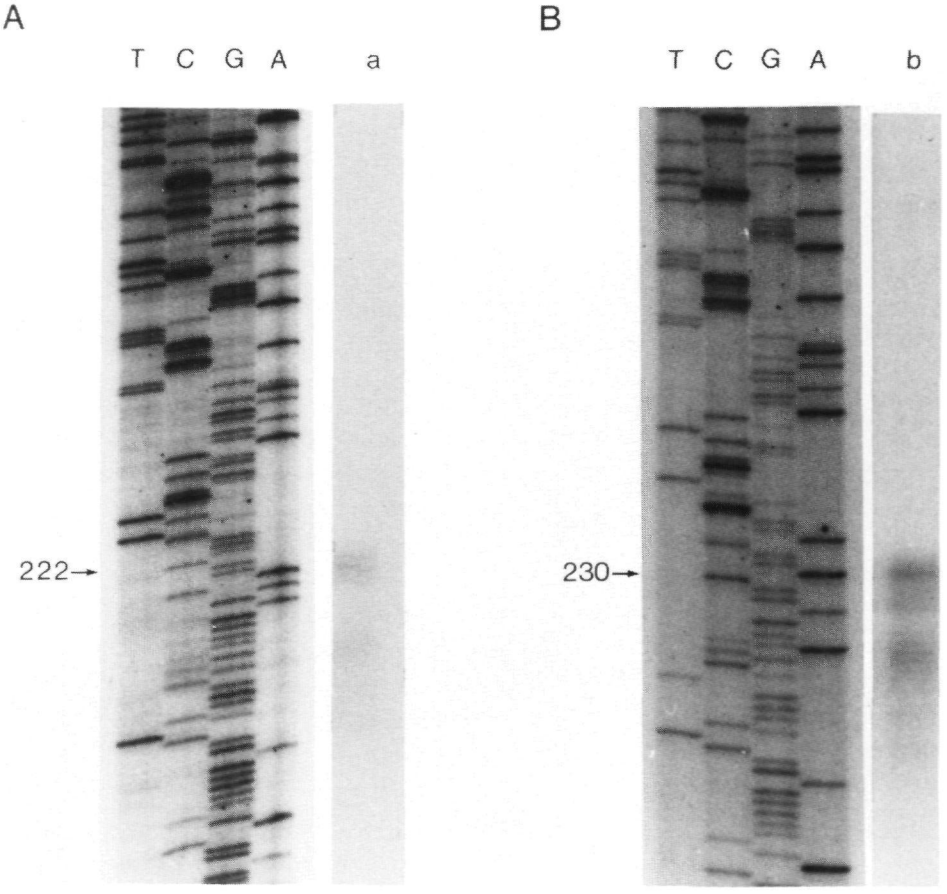
Screening of the human ML-1 cDNA library with the insert of pAO73 as a molecular probe resulted in the isolation of pHN1. To map the poly(A)-addition site in the human sis transcript, the nucleotide sequence of the 0.6 kbp insert of pHN1 was determined (data not shown). This sequence appeared to be identical to nucleotide sequences of cDNA clone pSM-1 described by Ratner et al. (15). The sequence of pSD1, a human sis-specific cDNA clone isolated by Rao et al. (16) lacked sequences at the 3' end. Comparison of the nucleotide sequence data of our human and feline sis cDNA clones indicated that the DNA region in which the poly(A)-addition site was located was well conserved. Only two small deletions in poly(A)-stretches (4 adenine residues in human and 1 in feline; see Fig. 1) were observed. In both

species, the same nucleotide constituted the poly(A)-addition site and the same site was also found by Ratner et al. (15). The poly(A)-addition site proposed by Rao et al. (16) can be found 30 nucleotides upstream of the site proposed in this paper (Fig. 1 and 2). No authentic polyadenylation signal (AATAAA) was observed in our isolates of the feline and human c-sis proto-oncogene. Rao et al. (16) have described an AATAAA sequence 18 bp downstream of their poly(A)-addition site. In our feline and human cDNA clones, a "C" instead of a "T" was present in the sequence at that position which does not represent the consensus sequence of the polyadenylation signal (Fig. 1; Fig. 2· nucleotide position 3347 to 3352).

Recently, it was suggested that the pentanucleotide sequence AUUUA in the 3′ region of mRNAs is involved in mRNA degradation (33). This sequence was mainly found in mRNAs that encode growth factors such as lymphokines and cytokines (33). The sequence can be present in multiple copies. For instance, human interleukin-3 contains 6 copies (L. Dorssers, pers. comm.). The 3′ region of the feline c-sis transcript appeared to contain 3 copies of this sequence (Fig. 2: nucleotide positions 2997, 3301 and 3314). The direct involvement of these sequences in sis mRNA degradation remains to be established.

Identification of the cap site.

To define precisely the 5′ end of the human and feline sis transcripts, nuclease S1 protection experiments were performed using mRNA from specimens of feline term placenta or human PC3 mRNA. A 235 bp AluI/AluI DNA fragment (A53; Fig. 2: nucleotide position -13 to 222 of feline c-sis) and a 243 bp AluI/AluI DNA fragment (A44; nucleotide position 391 to 633 of human c-sis (13)) were used in these studies. Fig. 4 shows the results of the nuclease S1 digestion of hybrids between feline placenta RNA and A53 DNA (purified insert) and between PC3 mRNA and A44 DNA (purified insert). In analysis of the feline sis transcript, a hybrid of 222 nucleotides long appeared to represent a major portion of the nuclease S1 resistant molecules (Fig. 4, lane a). Some other bands were visible on the autoradiographs and they represented hybrids of different length. Similar results were obtained in nuclease S1 experiments with the human sis transcript (Fig. 4, lane b). The results suggest that the feline and human sis transcript have their cap sites in the same position. The presence of some bands of minor intensity on the autoradiographs could be explained as follows. It is possible, that more transcription start sites are present. This is more often observed, as for instance with transcripts of the genes coding for chicken α2 type I col-



4. Nuclease S1 analysis of the feline and human sis transcripts. (A) Uniformly ^{32}P-labeled insert of A53 was hybridized with feline placenta mRNA. Upon treatment of the hybrid with nuclease S1, the product was analyzed under denaturing conditions (lane a). The four lanes at left show a portion of the nucleotide sequence of A53 to estimate molecular weights. (B) Nuclease S1 resistant hybrid products of human PC3 mRNA and A44. A portion of nucleotide sequence of A44 is shown at the left.

lagen, epidermal growth factor receptor and γ-crystallins (29,34,35). It is also possible that minor fractions of the double stranded hybrids are not properly digested by nuclease S1.

Based upon our data, the cap site of the sis transcript could be placed at the adenine residue at the nucleotide position 31 bp downstream of the "TATA" box. It was reported that most known eukaryotic mRNAs start with an

```
                                                                                ↓
Fe c-sis   MNRCWALFLSLCCYLRLVSAEGDPIPEELYKMLSDHSIRSFDDLQRLLHGDSVDEDRAELDLNSTRSHCGGELESLSRGRRSLGEAAGSPTVAE
Hu c-sis   ------------------------------E-------------------PGE--G------M----S-------A------    --L-I--
   v-sis                  MTLTWQ-------------G-------------Q---GK--G------M----S-------A--K---    --LS---

Fe c-sis   PAMIAECKTRTEVFEVSRRLIDRTNANFLVWPPCVEVQRCSGCCNNRNVQCRPTQVQLRLVQVRKIEIVRKRPVFKKATVTLEDHLACKCETVV
Hu c-sis   ---------------I-----------------------------------------------P----------K-I------------------A
   v-sis   ---------------I-----------------------------------------------P----------K-I-----------------I-A

Fe c-sis   AARPVTRSPGSSQEQRARTPQTRVTIRTVRVRRPPKGKHQKFKHTHDKKALKETLGA
Hu c-sis   ----------G------K--------------------R--------T--------
   v-sis   ---A------T------K-T-S----------------R-C------T--------
```

5. Comparison of the deduced amino acid sequences of the putative translational products of feline and human (14) c-sis and of v-sis of SSV (37). The arrow (↓) indicates the position of the N-terminus of the human PDGF-2 sequence (38).

adenine residue (36). (The adenine residue at nucleotide position 404 of the human c-sis proto-oncogene was mistakenly given as a "C" in a previous publication (13)). Therefore, our results are in good agreement with consensus characteristics of a cap site. Recently, Rao et al. (16) mapped the cap site of the human c-sis transcript at nucleotide position corresponding to position -6 of the feline DNA (Fig. 2). Messenger RNA isolated from human placenta and EJ tumor cells were used. In both cases, the same cap site was found and was located 25 bp downstream of the "TATA" box. The sequence of this cap site and its distance to the "TATA" box are not in good agreement with consensus cap site characteristics described by Breathnach and Chambon (36).

DNA insertion in feline c-sis.

Comparative nucleotide sequence analysis of genomic clones of feline and human c-sis, of cDNA clones of the sis transcripts of the two species and of the proviral DNA of the oncogene of SSV revealed two interesting differences. First of all, an insertion of 12 bp was found at the 3' end of exon 3 (Fig. 2: nucleotide position 1243 to 1254) of the feline c-sis gene. The ATG start codon is located at position 993 and the stop codon at position 1704. Therefore, the insert is present in the region that encodes the PDGF domain. To illustrate the effect of the insertion at the protein level, a comparison of the putative translational products of the feline, human and viral sis genes is shown in Fig. 5. The very high degree of homology is clear. The amino-terminus of the human PDGF-2 domain in the precursor is given by an arrow. In case the same proteolytic cleavage site is used in the feline and the human precursor protein, the 12 bp insertion could provide feline PDGF-2 with a different NH_2-terminus. Because of the nature of the amino acid sequence of the insertion, the different NH_2-terminus is not very likely to provide feline PDGF-2 with different characteristics.

It should be noted that the 5′ end of the 12 bp insertion resembles the consensus sequence of a 3′ splice junction. From analysis of our feline sis-specific cDNAs it can be concluded that this splice site is not used during splicing of feline sis precursor RNA. Theoretically, however, it is possible that the cDNAs described in this paper represent only a not fully spliced sis precursor RNA. The fact that no cDNAs without the insertion were isolated argues against the latter possibility.

When compared to the v-sis sequences of SSV, an insertion was found at the 5′ end of exon 7 of the feline c-sis proto-oncogene (Fig. 2: nucleotide position 1753 to 1895). This insertion was also present in the human c-sis proto-oncogene. It has been suggested (15) that because of differences in splicing of the human and woolly monkey sis precursor RNA, the sequences described above are deleted from the woolly monkey sis mRNA. Based upon the presence of this region in feline sis-specific cDNA clones, it is more likely that a deletion in the SSV genome occurred during or subsequently to its generation.

From the structure of the human and feline sis transcripts described in this paper, a consensus structure for the sis transcription unit can be deduced. The data strongly support our assignment of the promoter region that controls sis expression.

ACKNOWLEDGEMENTS

We thank Adrie van Bokhoven, Inge van Renesse, Peter Hendriksen and Saskya Ebeling for their excellent technical assistance. This work was supported by the Netherlands Cancer Foundation, Contract No. NUKC Bioch. 85-9.

REFERENCES

1. Doolittle, R.F., Hunkapiller, M.W., Hood, L.E., Devare, S.G., Robbins, K.C., Aaronson, S.A. and Antoniades, H.N. (1983) Science 221, 275-277.
2. Waterfield, M.D., Scrace, G.T., Whittle, N., Stroobant, P., Johnsson, A., Wasteson, A., Westermark, B., Heldin, C-H., Huang, J.S. and Deuel, T.F. (1983) Nature 304, 35-39.
3. Robbins, K.C., Antoniades, H.N., Devare, S.G., Hunkapiller, M.W. and Aaronson, S.A. (1983) Nature 305, 605-608.
4. Irgens, K., Wyers, M., Moraillon, A., Parodi, A. and Fortuny, V. (1973) C. R. Acad. Sci. Paris 26, 1783-1786.
5. Besmer, P., Snyder, H.W.,Jr., Murphy, J.E., Hardy, W.D.,Jr. and Parodi, A. (1983) J. Virol. 46, 606-613.
6 Hannink, M and Donoghue, D.J. (1984) Science 226, 1197-1199.
7. King, C.R., Giese, N.A., Robbins, K.C. and Aaronson, S.A. (1985) Proc. Natl. Acad. Sci. USA 82, 5295-5299.
8. Kelly, J.D., Raines, E.W., Ross, R. and Murray, M.J. (1985) EMBO J. 4, 3399-3405.
9. Chiu, I-M., Reddy, E.P., Givol, D., Robbins, K.C., Tronick, S.R. and Aaronson, S.A. (1984) Cell 37, 123-129.
10. Josephs, S.F., Guo, C., Ratner, L. and Wong-Staal, F. (1984) Science 223, 487-491.

11. Johnsson, A., Heldin, C-H., Wasteson, A., Westermark, B., Deuel, T.F., Huang, J.S., Seeburg, P.H., Gray, A., Ullrich, A., Scrace, G., Stroobant, P. and Waterfield, M.D. (1984) EMBO J. 3, 921-928.
12. Van den Ouweland, A.M.W., Breuer, M.L., Steenbergh, P.H., Schalken, J.A., Bloemers, H.P.J. and Van de Ven, W.J.M. (1985) Biochim. Biophys. Acta 825, 140-147.
13. Van den Ouweland, A.M.W., Roebroek, A.J.M., Schalken, J.A., Claesen, C.A.A., Bloemers, H.P.J. and Van de Ven, W.J.M. (1986) Nucl. Acids Res. 14, 765-778.
14. Collins, T., Ginsburg, D., Boss, J.M., Orkin, S.H. and Pober, J.S. (1985) Nature 316, 748-750.
15. Ratner, L., Josephs, S.F., Jarrett, R., Reitz, M.S.,Jr. and Wong-Staal, F. (1985) Nucl. Acids Res. 13, 5007-5018.
16. Rao, C.D., Igarashi, H., Chiu, I-M., Robbins, K.C. and Aaronson, S.A. (1986) Proc. Natl. Acad. Sci. USA 83, 2392-2396.
17. Takeda, K., Minowada, J. and Bloch, A. (1981) Cancer Res. 42, 5152-5157.
18. Dodemont, H.J., Soriano, P., Quax, W.J., Ramaekers, F., Lenstra, J.A., Groenen, M.A.M., Bernardi, G. and Bloemendal, H. (1982) EMBO J. 1, 167-171.
19. Messing, J. and Vieira, J. (1982) Gene 19, 269-276.
20. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
21. Staden, R. (1982) Nucl. Acids Res. 10, 2951-2961.
22. Auffray, C. and Rougeon, F. (1980) Eur. J. Biochem. 107, 303-314.
23. McMaster, G.K. and Carmichael, G.G. (1977) Proc. Natl. Acad. Sci. USA 74, 4835-4838.
24. Schalken, J.A., Van den Ouweland, A.M.W., Bloemers, H.P.J. and Van de Ven, W.J.M. (1985) Biochim. Biophys. Acta 824, 104-112.
25. Church, G.M. and Gilbert, W. (1984) Proc. Natl. Acad. Sci. USA 81, 1991-1995.
26. Roebroek, A.J.M., Schalken, J.A., Bussemakers, M.J.G., Van Heerikhuizen, H., Onnekink, C., Debruyne, F.M.J., Bloemers, H.P.J. and Van de Ven, W.J.M. (1986) Molec. Biol. Rep. 11, 117-125.
27. Huynh, T.V., Young, R.A. and Davis, R.W. (1985) in DNA cloning techniques, a practical approach (Glover, D. ed), IRL Press, Oxford.
28. Hanahan, D. and Meselson, M. (1980) Gene 10, 63-67.
29. Van Leen, R.W., Kastrop, P.M.M., Van Roozendaal, K.E.P. and Schoenmakers, J.G.G. (1986) Eur. J. Biochem. 157, 203-208.
30. Goustin, A.S., Betsholtz, C., Pfeifer-Ohlsson, S., Persson, H., Rydnert, J., Bywater, M., Holmgren, G., Heldin, C-H., Westermark, B. and Ohlsson, R. (1985) Cell 41, 301-312.
31. Slamon, D.J. and Cline, M.J. (1984) Proc. Natl. Acad. Sci. USA 81, 7141-7145.
32. Yarden, Y., Escobedo, J.A., Kuang, W-J., Yang-Feng, T.L., Daniel, T.O., Tremble, P.M., Chen, E.Y., Ando, M.E., Harkins, R.N., Francke, U., Fried, V.A., Ullrich, A. and Williams, L.T. (1986) Nature 323, 226-232.
33. Shaw, G. and Kamen, R. (1986) Cell 46, 659-667.
34. Vogeli, G., Ohkubo, H., Sobel, M.E., Yamada, Y., Pastan, I. and De Crombrugghe, B. (1981) Proc. Natl. Acad. Sci. USA 78, 5334-5338.
35. Ishii, S., Xu, Y-H., Stratton, R.H., Roe, B.A., Merlino, G.T. and Pastan, I. (1985) Proc. Natl. Acad. Sci. USA 82, 4920-4924.
36. Breathnach, R and Chambon, P (1981) Ann. Rev. Biochem. 50, 349-383.
37. Devare, S.G., Reddy, E.P., Law, J.D., Robbins, K.C. and Aaronson, S.A. (1983) Proc. Natl. Acad. Sci. USA 80, 731-735.
38. Betsholtz, C., Johnsson, A., Heldin, C-H., Westermark, B., Lind, P., Urdea, M.S., Eddy, R., Shows, T.B., Philpott, K., Mellor, A.L., Knott, T.J. and Scott, J. (1986) Nature 320, 695-699.

SUMMARY

SAMENVATTING

SUMMARY

With the discovery of oncogenes in the genome of acutely transforming RNA tumorviruses, more insight in the process of cancer was obtained. Based upon the function of the products of the viral oncogenes their key role in the induction of neoplasmas was established. The viral oncogenes show homology with cellular genes, called proto-onocogenes, present in the genome of probably all eukaryotes.

At this moment the function of most proto-oncogenes is not well established. However, more data become available leading to the conclusion that these genes are not only involved in tumorigenesis but also in normal cell growth and differentiation.

In chapter 1, a summary is given of some cellular processes following the binding of a growth factor to its receptor. This summary is given because one chain of platelet-derived growth factor (PDGF-2) is encoded by the sis proto-oncogene. This proto-oncogene is the topic in this thesis. At the end of this chapter some general (molecular) biological aspects of PDGF and the sis proto-oncogene are presented.

In chapter 2, a description of the isolation and characterization of the human and feline c-sis proto-oncogenes is given. Screening of human and feline cosmid libraries with a v-sis specific probe resulted in the isolation of several human and feline c-sis containing cosmid clones which led to the isolation of the complete c-sis sequences. By comparative analysis of the human and feline cosmid clones the identification was achieved of a new 5′ c-sis exon (exon 1) which was not present in the viral sis oncogene.

In chapter 3, the nucleotide sequence of exon 1 of the human and feline c-sis proto-oncogenes and their flanking sequences was studied. A promoter region ("TATA" box) and consensus sequences for Sp1 binding sites were found in similar positions in the c-sis proto-oncogenes of both species. It appeared that the sequences of exon 1 were conserved and included the ATG startcodon.

Regions upstream and downstream of the c-sis proto-oncogenes appeared also to be conserved. They contain possibly repetitive sequences which, however, do not belong to the category of highly repetitive sequences.

Whereas PDGF consists of two polypeptide chains, only one chain (PDGF-2) was identified in the c-sis containing cosmid clones, since a synthetic mixed oligonucleotide probe for PDGF-1 failed to detect homologous sequences.

In chapter 4, the complete nucleotide sequence of the feline c-sis exons is presented. Using nuclease S1 assays and cDNA cloning experiments, the cap site and poly(A)-addition site of the human and feline c-sis proto-oncogenes were determined. The sis transcripts of these two species, which are similar in size (3.5 kb), show a strong conservation of their 5′ and 3′ ends.

Comparative analysis of the deduced translational products of the human, feline and viral sis genes leads to the conclusion that these products are highly conserved. However, at the amino-terminal end of the feline PDGF-2 an insertion of 4 amino acids is present. Because of the nature of these amino acids it is not likely that feline PDGF has very different characteristics than human PDGF.

Compared to the v-sis sequences of simian sarcoma virus

(SSV), the human and feline c-sis proto-oncogenes contain an insertion at the 5′ end of exon 7. This deletion in the SSV genome is likely occurred during or subsequently to the generation of SSV.

Expression of c-sis is observed in a number of cell types and, in some cases, expression is modulated during cell differentiation. The biological reagents that are developped in the research project described in this thesis make it possible to study this modulation in c-sis expression and to define the nucleotide sequences that control these processes.

SAMENVATTING

Met de ontdekking van de aanwezigheid van oncogenen in het genoom van acuut transformerende RNA tumor virussen werd een grote stap voorwaarts gemaakt in het verkrijgen van inzicht in het ontstaan van kanker. De eiwitprodukten van de oncogenen blijken een sleutelrol te spelen in de inductie van neoplasmas. Deze virale oncogenen vertonen homologie met cellulaire genen, zogenaamde proto-oncogenen, aanwezig in het genoom van waarschijnlijk alle eukaryoten.

De funktie van de meeste proto-oncogenen is nog niet precies bekend. Er zijn echter sterke aanwijzingen dat deze genen betrokken zijn bij zowel de vorming van tumoren als in het groei- en differentiatieproces.

In hoofdstuk 1 wordt een overzicht gegeven van enkele processen in de cel die volgen op de binding van een groeifaktor aan zijn receptor. Dit overzicht is hier opgenomen omdat een van de ketens van platelet-derived growth factor (PDGF-2) gecodeerd wordt door het sis proto-oncogen, het gen dat in dit proefschrift centraal staat. Tot slot worden enkele algemene (moleculair) biologische aspecten van PDGF en het sis proto-oncogen gepresenteerd.

In hoofdstuk 2 wordt de isolatie en karakterisatie van het humane en katte c-sis proto-oncogen beschreven. Door gebruik te maken van cosmide banken, welke gemaakt zijn met behulp van humaan en katte chromosomaal DNA, en een v-sis specifieke probe, werden diverse c-sis bevattende cosmide clonen geïsoleerd. Dit resulteerde in de clonering van de totale humane en katte c-sis proto-oncogen sekwenties. Vergelijkende

studies van deze cosmide clonen leidden tot de beschrijving van een nieuw 5′ gelegen exon (=exon 1), dat niet aanwezig was in het virale sis oncogen.

In hoofdstuk 3 worden de sekwenties van exon 1 en zijn omgeving getoond zowel van het humane als het katte c-sis proto-oncogen. Een "TATA" box en consensus sekwenties voor Sp1 bindingsplaatsen zijn in het humane en in het katte c-sis proto-oncogen op identieke plaatsen gelegen. Verder bleek uit deze studie dat exon 1 sterk geconserveerd was en dat het startcodon voor de translatie in dit exon gelegen was.

Ook gebieden buiten het c-sis proto-oncogen bleken tussen mens en kat sterk geconserveerd te zijn. Mogelijk bevatten zij repeterende sekwenties. Deze behoren echter niet tot de klasse van vaak voorkomende repeterende sekwenties.

Door gebruik te maken van een gemengde synthetische oligonucleotide probe werd aangetoond dat in de humane c-sis bevattende cosmide clonen geen sekwenties aanwezig waren die zouden kunnen coderen voor de andere keten van PDGF (PDGF-1).

In hoofdstuk 4 wordt de volledige exon sekwentie van het katte c-sis proto-oncogen gepresenteerd. Met behulp van nuclease S1 assays en cDNA cloneringen werden respectievelijk de zgn. cap site en poly(A)-additie site in zowel het humane als katte c-sis proto-oncogen geplaatst. Hieruit bleek dat zowel het 5′ als 3′ uiteinde van het c-sis gen sterk geconserveerd waren tussen deze species. Het geconserveerde karakter van het c-sis gen kwam ook tot uitdrukking in de grootte van het sis mRNA: 3.5 kb voor zowel het humane als katte mRNA.

Uit de vergelijking van de theoretische translatie produkten van het humane, katte en virale sis gen volgde de conclusie dat

ook de eiwitprodukten geconserveerd zijn. Echter, aan het amino-terminale uiteinde van PDGF-2 is bij de kat een insertie van 4 aminozuren gevonden. Vanwege de aard van deze extra aminozuren lijkt het niet waarschijnlijk dat PDGF van de kat sterk verschillende eigenschappen zou hebben dan humaan PDGF.

Uit de vergelijking van de DNA sekwenties van de humane en katte c-sis genen enerzijds en het virale sis gen anderzijds werd duidelijk dat aan het 5′ einde van exon 7 bij zowel het humane als het katte gen extra sekwenties aanwezig waren. Hoogstwaarschijnlijk zijn deze sekwenties gedurende de vorming van het simian sarcoma virus, dat het v-sis oncogen bevat, verloren gegaan.

Uit deze studie naar het c-sis proto-oncogen blijkt dat dit gen tussen mens en kat sterk geconserveerd is, zelfs buiten de exon sekwenties. Met behulp van de gegevens in dit proefschrift kunnen verdere studies naar de regulatie van de expressie van het sis proto-oncogen verricht worden.

LIST OF PUBLICATIONS

- Anton Berns, Ans van den Ouweland, Wim Quint, Jan van Oirschot and Arno Gielkens (1985) Presence of markers for virulence in the unique short region or repeat region or both of Pseudorabies Hybrid Viruses. J. Virol. 53, 89-93.

- Joseph S. Verbeek, Anton J.M. Roebroek, Ans M.W. van den Ouweland, Henri P.J. Bloemers and Wim J.M. Van de Ven (1985) Human c-*fms* proto-oncogene: comparative analysis with an abnormal allele. Mol. Cell. Biol. 5, 422-426.

- Jack A. Schalken, Ans M.W. van den Ouweland, Henri P.J. Bloemers and Wim J.M. van de Ven (1985) Characterization of the feline c-*abl* proto-oncogene. Biochim. Biophys. Acta 824, 104-112.

- Joseph, S. Verbeek, Ans M.W. van den Ouweland, Jack A. Schalken, Anton, J.M. Roebroek, Carla Onnekink, Henri P.J. Bloemers and Wim J.M. van de Ven (1985) Molecular cloning of the feline c-*fes* proto-oncogene and construction of a chimeric transforming gene. Gene 35, 33-43.

- Ans M.W. van den Ouweland, Marco L. Breuer, Paul H. Steenbergh, Jack A. Schalken, Henri P.J. Bloemers and Wim J.M. Van de Ven (1985) Comparative analysis of the human and feline c-*sis* proto-oncogenes. Identification of 5′ human c-*sis* coding sequences that are not homologous to the transforming gene of simian sarcoma virus. Biochim. Biophys. Acta 825,

140-147.

- Anton J.M. Roebroek, Jack A. Schalken, Joseph S. Verbeek, Ans M.W. Van den Ouweland, Carla Onnekink, Henri P.J. Bloemers and Wim J.M. Van de Ven (1985) The structure of the human c-fes/fps proto-oncogene. EMBO J. 4, 2897-2903.

- Ans M.W. Van den Ouweland, Anton J.M. Roebroek, Jack A. Schalken, Chris A.A. Claesen, Henri P.J. Bloemers and Wim J.M. Van de Ven (1986) Structure and nucleotide sequence of the 5′ region of the human and feline c-sis proto-oncogenes. Nucl. Acids Res. 14, 765-778.

- Ans M.W. van den Ouweland, Jan J.M. Van Groningen, Jack A. Schalken, Han, W. Van Neck, Henri P.J. Bloemers and Wim J.M. Van de Ven (1987) Genetic organization of the c-sis transcription unit. Nucl. Acids Res. 15, 959-970.

CURRICULUM VITAE

Ans van den Ouweland werd op 28 november 1955 geboren te Tilburg. Na het behalen van het Atheneum B diploma (1974) en het diploma Biochemisch analiste aan de HLMS te Oss (1977) werd gestart met de studie Biologie aan de Landbouw Universiteit te Wageningen.

In 1983 werd het doctoraalexamen met goed gevolg afgelegd (cum laude) met als hoofdvakken erfelijkheidsleer (Prof. Dr. Ir. J.H. van der Veen ; o.l.v. Dr. Ir. J. Visser), experimentele diermorfologie en celbiologie (Prof. Dr. W.B. van Muiswinkel) en moleculaire biologie (Prof. Dr. A. van Kammen; o.l.v. Dr. A.J.M. Bisseling) en een stage gedurende 6 maanden op de afdeling Biochemie (Prof. Dr. H. Bloemendal; o.l.v. Dr. A.J.M. Berns, Katholieke Universiteit te Nijmegen).

Vanaf februari 1983 was zij verbonden als wetenschappelijk assistent aan de afdeling Biochemie van de Katholieke Universiteit te Nijmegen. Gedurende deze periode werd onderzoek verricht aan de humane en katte c-sis proto-oncogenen in de werkgroep o.l.v. Prof. Dr. H.P.J. Bloemers en Dr. W.J.M. van de Ven.

Vanaf februari 1987 is zij op bovengenoemde afdeling werkzaam op een door het Koningin Wilhelmina Fonds gesubsidieerd projekt dat de verdere karakterisatie van het fur gen tot doel heeft.

STELLINGEN

1. De conclusie dat epidermale cellen gevoeliger zijn voor UVA licht en een psoralen/UVA behandeling kan niet getrokken worden uit een vergelijking van de resultaten verkregen met een muize-keratinocyten-cellijn met die van een menselijke epitheliale carcinoma-cellijn.

   Laskin et al. (1986) Proc. Natl. Acad. Sci. USA 83, 8211-8215

2. De onzekerheid bij Rao et al. inzake de positie van de poly(A)-additie plaats in het menselijke c-sis proto-oncogen is te wijten aan het niet optimaal gebruik maken van eigen resultaten en reeds gepubliceerde gegevens.

   Rao et al. (1986) Proc. Natl. Acad. Sci. USA 83, 2392-2396.
   Ratner et al. (1985) Nucl. Acids Res. 13, 5007-5018.

3. Het gebruik van zgn. primer extension produkten, verkregen met RNA als matrijs, als moleculaire probes om RNA expressie te bestuderen is een hachelijke zaak.

4. Een duidelijk onderscheid tussen de beide vormen en de twee ketens van "platelet-derived growth factor" moet de grondslag zijn voor een éénduidige nomenclatuur.

   Tong et al. (1986) Mol. Cell. Biol. 6, 3018-3022.

5. Het gebruik van baculovirussen als expressiesysteem voor de produktie van eiwitten, uitgaande van cDNA klonen, voor de bereiding van monoclonale antilichamen heeft de voorkeur boven een bacteriëel expressiesysteem.

6. Door gebruik te maken van een antiserum dat zowel het α-crystalline als enkele andere eiwitten herkent wordt veel afbreuk gedaan aan de conclusie dat Müller-glia-cellen mogelijk α-crystalline bevatten.

   Moscona et al. (1985) Proc. Natl. Acad. Sci. USA 82, 5570-5573.

7. De suggestie dat de 5'-sekwenties van het BCR-ABL mRNA identiek zijn aan sekwenties aanwezig in het 4.5 kb BCR mRNA is nogal voorbarig.

   Mes-Masson *et al*. (1986) Proc. Natl. Acad. Sci. USA *83*, 9768-9772.

8. Door de koppeling van het alarmsysteem aan de "aan/uit" knop bij de Bosch diepvrieskast (type GSL 3000) wordt bij storingen in het elektriciteitscircuit geen alarm gegeven zodat geen optimaal rendement van dit systeem verkregen wordt.

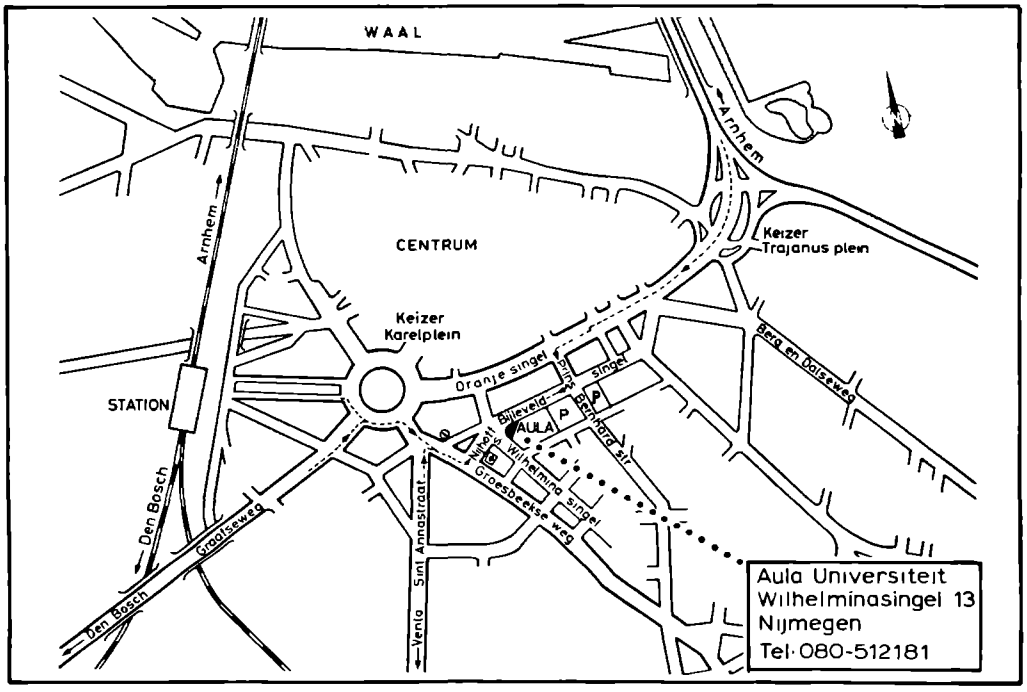
9. Een meer op prestatie gericht personeelsbeleid zal er toe bijdragen de uittocht van universitair personeel naar het bedrijfsleven te verminderen.

10. De versoepeling van de volleybal-technische regels is voor veel NeVoBo scheidsrechters een aanleiding om hun licenties naar de NeVoBo terug te sturen.

Stellingen behorend bij het proefschrift:

"Characterization of the human and feline c-*sis* proto-oncogenes".

12 juni 1987 Ans van den Ouweland


WAAL
Arnhem
CENTRUM
Keizer
Trajanus plein
Keizer
Karelplein
Oranje singel
Prins
Singel
Berg en Dalseweg
STATION
Billeveld
AULA
P
Reinaerd str
Wilhelmina
Wilhelmina singel
Groesbeekse weg
Den Bosch
Graalseweg
Sint Annastraat
Venlo
Aula Universiteit
Wilhelminasingel 13
Nijmegen
Tel·080-512181

RECEPTIE NA AFLOOP VAN DE PROMOTIE
IN DE AULA
VAN DE KATHOLIEKE UNIVERSITEIT,
WILHELMINASINGEL 13,
NIJMEGEN

Printed by Modern, Bennekom